Optimal Taxation of Robots

Uwe Thuemmel*
University of Zurich†

August 2018

Abstract

I study the optimal taxation of robots and labor income. In the model, robots substitute for routine labor and complement non-routine labor. I show that while it is optimal to distort the use of robots, robots may be either taxed or subsidized. The robot tax exploits general-equilibrium effects to compress the wage distribution. Wage compression reduces income-tax distortions of labor supply, thereby raising welfare. Quantitatively, the optimal robot tax equals 4% if occupations are fixed, but its welfare impact is negligible. With occupational choice, the optimal robot tax is 0.4%, and approaches zero as the price of robots falls.

Key words: Optimal taxation, Input taxation, Production efficiency, Technological change, Robots, Inequality, General equilibrium, Multidimensional heterogeneity

JEL-Codes: D31, D33, D50, H21, H23, H24, H25, J24, J31, O33

---

*I am especially grateful to Florian Scheuer, Bas Jacobs, and Bjoern Bruegemann for very helpful conversations and comments. I also thank Eric Bartelsman, Michael Best, Nicholas Bloom, Raj Chetty, Iulian Ciobica, David Dorn, Aart Gerritsen, Renato Gomes, Jonathan Heathcote, Gee Hee Hong, Hugo Hopenhayn, Caroline Hoxby, Albert Jan Hummel, Guido Imbens, Pete Klenow, Wojciech Kopczuk, Tom Krebs, Dirk Krueger, Per Krusell, Simas Kucinskas, Sang Yoon Lee, Moritz Lenel, Agnieszka Markiewicz, Magne Mogstad, Serdar Ozkan, Pascual Restrepo, Dominik Sachs, Michael Saunders, Isaac Sorkin, Kevin Spiritus, Nicolas Werquin, Frank Wolak, Floris Zoutman as well as seminar participants at Erasmus University Rotterdam, VU Amsterdam, CPB Netherlands, and participants of the CESifo Public Economics Area Conference for valuable comments and suggestions. This paper has benefited from a research visit at Stanford University supported by the C. Willems Stichting. All errors are my own.

†Address: University of Zurich, Department of Economics, Schoenberggasse 1, 8001 Zurich, Switzerland. E-mail: uwe.thuemmel@uzh.ch. Web: http://uwethuemmel.com/
1 Introduction

Public concern about the distributional consequences of automation is growing (see e.g. Ford, 2015; Brynjolfsson and McAfee, 2014; Frey et al., 2017). It is feared that the “rise of the robots” is going to disrupt the labor market and will lead to extreme income inequality. These concerns have raised the question how redistributive policy should respond to automation. Some policy makers and opinion leaders have suggested a “tax on robots”. Is this a good idea? This paper tries to answer that question. I find that while it is generally optimal to distort the use of robots, robots may be either taxed or subsidized. Quantitatively, the optimal robot tax is positive, but its welfare impact is negligible.

The robot tax exploits general equilibrium effects to compress the wage distribution. Wage compression makes it less distortionary to tax income, which allows for more redistribution overall and raises welfare. If robots primarily substitute for routine labor at medium incomes, a tax on robots decreases wage inequality at the top of the wage distribution, but raises inequality at the bottom. The sign of the robot tax is then theoretically ambiguous. Quantitatively, in the United States, the optimal robot tax is about 4% in the short-run with fixed occupations, but is diminished to 0.4% in the medium-run with occupational choice – approaching zero as robots get cheaper. Both in the short- and medium-run, the welfare impact of introducing a robot tax is negligible.

To reach these conclusions, I first build intuition by studying a stylized model based on Stiglitz (1982). The full model then embeds labor market polarization similar to Autor and Dorn (2013) in an optimal taxation framework based on Rothschild and Scheuer (2013, 2014). For the quantitative analysis, I calibrate the full model to the US economy, using data from the Current Population Survey (CPS), as well as evidence on the impact of robots on the labor market from Acemoglu and Restrepo (2017).

The stylized model extends Stiglitz (1982) to three occupations: non-routine manual, routine, and non-routine cognitive. It captures that cognitive non-routine workers earn on average higher wages than routine workers, who in turn earn higher wages than manual non-routine workers (see e.g. Acemoglu and Autor, 2011). Workers are fixed-assigned to one of the three occupations. Moreover, while in Stiglitz (1982) output is produced by labor only, I introduce robots as additional production factor. Crucially, robots are more complementary to non-routine labor than to routine labor. An increase in the amount of robots therefore lowers demand and wages for routine workers relative to non-routine workers. The robot tax exploits this differential impact of robots on wages. I derive a formula for the optimal robot tax which features elasticities of relative wage rates with respect to robots and incentive effects. Under the realistic assumption that income taxes may not be conditioned on occupation, the robot tax is in general not zero, violating production efficiency (see Diamond and Mirrlees, 1971). However, it is theoretically ambiguous whether robots should be taxed or subsidized. This is because the robot tax has counteracting effects on wages at the top and the bottom of the wage distribution. Ceteris paribus, the robot tax is larger, the more robots raise wage inequality between non-routine cognitive workers and routine workers; it is lower, the more robots compress the wage gap between routine workers and manual non-routine workers. The incentive effects capture how much a change in relative wages affects income-tax distortions of labor supply. In addition,

---

1See e.g. this quote from a Draft report by the Committee on Legal Affairs of the European Parliament.

“Bearing in mind the effects that the development and deployment of robotics and AI might have on employment and, consequently, on the viability of the social security systems of the Member States, consideration should be given to the possible need to introduce corporate reporting requirements on the extent and proportion of the contribution of robotics and AI to the economic results of a company for the purpose of taxation and social security contributions;”

2Bill Gates has advocated for a tax on robots. See https://qz.com/911968/bill-gates-the-robot-that-takes-your-job-should-pay-taxes/
they capture how much the government values redistribution between workers with different incomes. Ceteris paribus: the more a reduction of inequality at the top of the wage distribution lowers income-tax distortions of labor supply, the larger is the optimal tax on robots; moreover, the robot tax is larger, the more the government cares about reducing income-tax distortions at the top. In contrast, the tax on robots is smaller, the more an expansion of inequality at the bottom of the wage distribution worsens income-tax distortions of labor supply, and the more the government cares about these.

What the sign and size of the optimal robot tax should be therefore turns into a quantitative question. The stylized model misses two features which are particularly important for a quantitative analysis: continuous wage distributions which overlap occupations, and the possibility of switching occupations. With continuous wages, the impact of robots on inequality can be captured more realistically, while occupational choice is a relevant margin of adjustment to automation. The full model incorporates both features, building upon Rothschild and Scheuer (2013). Individuals now differ in their three-dimensional ability, based on which they choose labor supply and their occupation. The expression for the optimal robot tax has a similar structure as in the stylized model: elasticities of relative wage rates with respect to robots are again central. In addition, a tax on robots now leads to a reallocation of labor supply within occupations, which affects how much labor supply is distorted by income taxation. In the robot tax formula, this is captured by effort reallocation effects. Also, since a tax on robots drives up wages in routine occupations relative to non-routine occupations, some non-routine workers find it beneficial to switch to routine work. The robot tax formula captures this by including occupational shift effects. By switching occupations, individuals offset part of the wage compression which can be achieved by taxing robots. As a result, the robot tax becomes a less effective policy instrument.

To assess the optimal policy quantitatively, I calibrate the full model to the US economy. To do so, I use data on wages and occupational choice from the CPS. Moreover, I calibrate the impact of robots on wages based on Acemoglu and Restrepo (2017). I compute optimal policy for two scenarios: in the short-run, occupations are fixed, while in the medium-run, occupational switching becomes possible. The short-run optimal robot tax is around 4%. The welfare gain of introducing a robot tax expressed in dollars per person per year is in the order of 0.80$. In the medium-run with occupational choice, the tax is in the order of 0.4%, and its welfare impact is reduced to 0.01$. Moreover, the robot tax and its welfare impact go to zero as the price of robots falls. This is a consequence of individuals moving out of routine occupations as robots substitute for them. With fewer and fewer individuals left in routine jobs, the welfare impact of raising their relative wages goes to zero. I conclude that in light of the negligible welfare effects, this paper does not provide a strong case for taxing robots.

The remainder of the paper is structured as follows: Section 2 discusses the related literature. Section 3 sets up a simplified model with discrete worker types and without occupational choice to build intuition. Section 4 introduces continuous types and occupational choice and characterizes the optimal robot tax in the full model. Section 5 studies the quantitative implications of the model. Section 6 discusses the results and concludes. Proofs and additional material are contained in an Appendix.

2 Related literature

This paper builds upon the framework by Rothschild and Scheuer (2013, 2014) who study optimal non-linear income taxation with multi-dimensional heterogeneity and sectoral choice, thereby extending and generalizing Stiglitz (1982). The modeling of the economy in the quantitative part combines a production technology similar to Autor and Dorn (2013) with a Roy (1951) model of occupational choice. In addition, this paper is related to different strands in the literature which are discussed below.
Optimal taxation and technological change. A growing number of papers investigates the question how taxes should respond to technological change. Most closely related is Guerreiro et al. (2017), who in parallel and independent work also ask whether robots should be taxed. Their model features two discrete types of workers – routine and non-routine – who are assigned tasks. In addition, some tasks are performed by robots. In a model like Stiglitz (1982), in which labor income is taxed non-linearly, they show that it is optimal to tax robots (provided that some tasks are still performed by routine labor). The rationale for taxing robots is the same as in this paper: compressing the wage distribution to reduce income-tax distortions of labor supply. Arguing that such a non-linear tax system can be complex and difficult to implement, Guerreiro et al. then focus on a parametric tax schedule: first, following Heathcote et al. (2017), and then augmenting the tax system with a lump-sum rebate. Under these tax systems, a tax on robots is also optimal. Finally, they introduce occupational choice by assuming that individuals have different preferences for non-routine work. As in this paper, the occupational choice margin matters for redistribution via general-equilibrium effects. In addition, like this paper, Guerreiro et al. find that the optimal robot tax is lower if occupational choice is possible (though they do not discuss this result). They conclude that, unless the current US tax system is reformed, a drop in the price of robots will lead to massive income inequality; and that, unless there is full automation, it is optimal to tax robots.

This paper differs from Guerreiro et al. (2017) in important ways. First, by writing technology as a function of the aggregate amount of robots, this paper arrives at optimal tax expressions which are easily interpretable. For example, they feature the elasticities of relative wage rates with respect to robots. Second, by considering three groups of occupations, I allow for wage polarization. The empirical literature on the labor market effects of technological change has highlighted that routine workers are found in the middle of the income distribution (see e.g. Acemoglu and Autor, 2011). The sign of the robot tax is then theoretically ambiguous. Third, my model features heterogeneity within occupations and thus generates a realistic wage distribution, while the model by Guerreiro et al. only features two levels of wages in the economy. Fourth, I model occupational choice based on individuals’ earnings abilities in different occupations, relating to the literature on employment polarization (see e.g. Acemoglu and Autor, 2011; Autor and Dorn, 2013; Goos et al., 2014; Cortes, 2016). As a consequence, wage distributions across occupations overlap, as observed in the data. Moreover, workers who move out of routine occupations in response to automation experience a drop in earnings relative to non-routine workers. In contrast, in the model by Guerreiro et al., workers who switch from routine to non-routine work see their incomes rise, which is counterfactual (see Cortes et al., 2017).

Eventually, what the optimal level of the robot tax should be is a quantitative question. Here, my analysis goes beyond that of Guerreiro et al. whose numerical examples are mostly illustrative. I calibrate my model based on data for the US economy and match the distribution of incomes and employment. Moreover, I use the empirical evidence on the labor market effects of robots by Acemoglu and Restrepo (2017). Based on the calibrated model, I find an optimal robot tax which is substantially lower than the maximum levels found by Guerreiro et al. Finally, while Guerreiro et al. compare welfare across the different tax systems, they do not isolate the welfare impact of introducing a robot tax – though this is arguably a relevant number to answer the question whether robots should be taxed. Based on the negligible welfare gains of taxing robots, this paper does not provide a strong case for taxing robots.

Related, Costinot and Werning (2018a,b) ask how tax policy should respond to inequality driven by technology or trade if the set of policy instruments is limited, such that production

---

3For example, in their simulations, routine and non-routine workers initially earn the same wage and make up equal shares of the population.

4They find an optimal robot tax of up to 10% in the model with non-linear taxes. In the version in which they augment the parametric tax function by Heathcote et al. (2017) with a lump-sum rebate, the robot tax reaches up to 30%. 

4
efficiency as in Diamond and Mirrlees (1971) is not optimal (see below for more on the relation to production efficiency). As one application, they study the optimal tax on robots, assuming that labor income can be taxed non-linearly, but may not depend on a worker’s type. As in this paper, if wages of different workers are differentially affected by robots, taxing robots is optimal to reduce inequality in order to dampen income-tax distortions of labor supply. For a general production technology, Costinot and Werning derive a sufficient statistics formula for the optimal robot tax which depends only on elasticities, factor shares, and marginal income tax rates. One important ingredient is the elasticity of wages with respect to robots, which is also central for optimal robot taxation in this paper.

The sufficient statistics approach makes the paper by Costinot and Werning (2018a,b) complementary to this paper. It allows to make statements about the optimal robot tax without having to assume a lot of structure on the economy. However, the sufficient statistics formula is only valid if the economy is already at a policy optimum. In contrast, the more structural approach in this paper does not impose that restriction. Moreover, it allows to analyze counterfactuals, such as the impact of a drop in the robot price on the optimal robot tax. In fact, Costinot and Werning (2018a) also assume more structure when they analyze the impact of a drop in the price of robots on the optimal robot tax in a stylized model. They show that despite robots being used more and inequality growing, the optimal robot tax falls, which they also demonstrate in a numerical example. In my fully calibrated quantitative analysis, I find as well that the optimal robot tax may drop as robots get cheaper.

Related, Tsyvinski and Werquin (2018) derive how a given tax system needs to be adjusted to compensate individuals for the distributional effects of, for example, trade or automation. In an application, they use the results from Acemoglu and Restrepo (2017) to investigate how individuals should be compensated for income losses and gains generated by the increased use of industrial robots. In contrast to this paper, Tsyvinski and Werquin (2018) do not study optimal taxation. Other papers (Gasteiger and Prettner, 2017; Hemous and Olsen, 2018) study the impact of taxing robots, taking a positive – rather than a normative – perspective.

The implications of technological change for tax policy are also analyzed by Ales et al. (2015) who study a model in which individuals are assigned to tasks based on comparative advantage. They ask how marginal tax rates should optimally have been set in the 2000s compared to the 1970s, based on changes in the US distribution of wages over a set of occupations. Over this period, labor market polarization has led to relative losses for middle income workers. As a consequence, the optimal income tax reform eases the burden for these workers. However, Ales et al. (2015) do not model automation, and thus do not study robot taxation. In my model, technological change – in this case robotization – also polarizes the income distribution, with consequences for the optimal income tax. In addition, a tax on robots has relative benefits for middle income workers and partly offsets wage polarization.

Production efficiency. A tax on robots violates production efficiency. This paper is thus related to the Production Efficiency Theorem (Diamond and Mirrlees, 1971) which states that production decisions should not be distorted, provided that the government can tax all production factors – inputs and outputs – linearly and at different rates. In addition, the Atkinson-Stiglitz Theorem (Atkinson and Stiglitz, 1976) states that if utility is weakly separable between consumption and leisure and the government can use a non-linear income tax, commodity taxes should not be used for redistribution. Combining the two theorems implies that neither consumption nor production should be distorted for redistributive reasons, provided the government can tax labor income non-linearly and has access to sufficient instruments to tax inputs and outputs. This implication has subsequently been put in perspective by Naito (1999); Saez (2004); Naito (2004); Jacobs (2015); Shourideh and Hosseini (2018); Costinot and Werning (2018a,b) who all study settings which feature fewer tax instruments than required for achieving production efficiency. Similarly, in this paper, the set of tax instruments is too
restricted for production efficiency to be optimal. In particular, income taxes may not be conditioned on occupation.\(^5\) In a related setting, Scheuer (2014) studies optimal taxation of labor income and entrepreneurial profits. He shows that when labor income and profits are subject to the same non-linear tax schedule, it is optimal to distort production efficiency in order to compress wages differentially. Production efficiency is restored if labor income and profits can be subject to different tax schedules.\(^6\) This paper focuses on the realistic case in which income taxes are not conditioned on occupation.

**Robots and the labor market.** A recent empirical literature studies the impact of robots on the labor market. Using data on industrial robots from the International Federation of Robotics (IFR, 2014), Acemoglu and Restrepo (2017) exploit variation in exposure to robots across US commuting zones to identify the causal effect of industrial robots on employment and wages between 1990 and 2007. I use their results to inform the quantitative analysis. Other articles which study the impact of robots on labor markets are Graetz and Michaels (2018) for a panel of 17 countries and Dauth et al. (2017) for Germany. A related literature studies labor market polarization due to technological change (see e.g. Goos et al., 2014; Cortes et al., 2017). Autor and Dorn (2013) investigate the impact of ICT technology on wages and employment in routine and non-routine occupations. In their model, ICT technology substitutes for routine work and complements non-routine work. Since qualitatively, robots can have a similar effect as ICT technology, I model technology in a related way.

**Taxation of capital.** Robots are a specific type of capital, which relates this paper to the literature on capital taxation. However, most arguments for taxing capital do not depend on the differential impact of capital on wages. Such arguments are therefore orthogonal to the reason for which robots are taxed in this paper. An exception is Slavík and Yazici (2014) who give a similar argument for taxing equipment capital as this paper does for taxing robots. Due to capital-skill complementarity (Krusell et al., 2000), a tax on equipment capital depresses the skill-premium, thereby reducing income-tax distortions of labor supply. In contrast, structures capital, which is equally complementary to low and high-skilled labor, should not be taxed. In their quantitative analysis for the US economy, they find an optimal tax on equipment capital of almost 40%. Moreover, they find large welfare gains of moving from non-differentiated to differentiated capital taxation. One reason for the different welfare implications is that I study the effect of introducing a robot tax into a system which taxes labor income optimally, whereas Slavík and Yazici start out from the current US tax system in which this is not the case. Moreover, I allow for robots to polarize the wage distribution, which leads to counteracting effects of the robot tax on the top and the bottom of the wage distribution – and thus to a lower tax rate.\(^7\)

### 3 Model with discrete types and no occupational choice

To develop intuition, I first discuss a simple model which features discrete types and abstracts from occupational choice. The model extends Stiglitz (1982) to three sectors (or occupations) and features endogenous wages. The model illustrates the key arguments for taxing robots. However, it is too stylized for a quantitative analysis, and by abstracting from occupational

\(^5\)Saez (2004) refers to this as a violation of the labor types observability assumption.

\(^6\)Gomes et al. (2018) set up a model in which workers with continuously distributed ability choose both, intensive margin labor supply and occupation, as they do in this paper. They then study optimal occupation-specific non-linear income taxation and show that occupational choice is optimally distorted. They refer to this as a distortion of production efficiency. In Scheuer (2014), occupational choice is distorted in the presence of occupation-specific taxes – however, production is efficient.

\(^7\)In Slavík and Yazici (2014) the returns to capital are taxed, whereas in my model the tax is levied on the stock of robots, which is another reason for the smaller magnitude of taxes in my paper.
choice leaves out an important adjustment margin. In Section 4, I discuss a richer model with continuous types and occupational choice which is amenable to a realistic calibration, and thus suitable to analyze the optimal taxation of robots quantitatively.

3.1 Setup

3.1.1 Workers, occupations and preferences

There are three types of workers $i \in I \equiv \{M, R, C\}$ with corresponding mass $f_i$. A worker’s type corresponds to his occupation, where $M$ refers to an occupation which requires manual non-routine labor, $R$ refers to an occupation requiring routine labor, and $C$ denotes a cognitive non-routine occupation. The distinction between routine and non-routine occupations is motivated by the empirical literature which has established that in recent decades technology has substituted for routine work, and has complemented non-routine work (see e.g. Autor et al., 2003). Moreover, the literature on labor market polarization suggests to distinguish between low-skilled and high-skilled non-routine occupations (see e.g. Cortes, 2016; Cortes et al., 2017).

Workers derive utility from consumption $c$ and disutility from labor supply $\ell$, according to the quasi-linear utility function

$$U(c, \ell) = c - \frac{\ell^{1+\frac{1}{\varepsilon}}}{1 + \frac{1}{\varepsilon}},$$

(1)

where $\varepsilon$ is the labor supply elasticity.

3.1.2 Technology

Denote by $L \equiv (L_M, L_R, L_C)$ the vector of aggregate labor supplies with $L_i = f_i\ell_i$ for all $i \in I$. Let $B$ denote robots. The final good is produced by a representative firm according to a constant returns to scale production function $Y(L, B)$. The firm maximizes profits by choosing the amount of total labor of each type $i \in I$ and the number of robots, taking wages $w_i$ and the price of robots $p$ as given. Normalizing the price of the final good to one, the firm’s profit maximization problem is

$$\max_{L, B} Y(L, B) - \sum_{i \in I} w_i L_i - pB.$$  

(2)

Denote the marginal products of total effective labor as

$$Y_i(L, B) \equiv \frac{\partial Y(L, B)}{\partial L_i} \forall i \in I,$$

(3)

and the marginal product with respect to robots as

$$Y_B(L, B) \equiv \frac{\partial Y(L, B)}{\partial B}.$$  

(4)

In equilibrium, we then have $w_i(L, B) = Y_i(L, B) \forall i \in I$ and $p = Y_B(L, B)$. Unless stated otherwise, I assume throughout that robots are better substitutes for routine work than for non-routine work. More specifically, I make the following assumption

**Assumption 1.** A marginal increase in the amount of robots raises the wage rate of non-routine labor relative to routine labor.

- $\frac{\partial}{\partial B} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) > 0,$
- $\frac{\partial}{\partial B} \left( \frac{Y_C(L, B)}{Y_R(L, B)} \right) > 0.$
Due to constant returns to scale, equilibrium profits are zero. Robots are produced linearly with the final good, according to

\[ B(x) = \frac{1}{q} x, \]  

(5)

where I denote by \( x \) the amount of the final good allocated to the production of robots, and where \( 1/q \) is the marginal rate of transformation between robots and the consumption good. In the absence of taxes, we then have \( p = q \) in equilibrium. Later, when taxes drive a wedge between \( p \) and \( q \), I refer to \( q \) as the producer price of robots and to \( p \) as the user price of robots.

3.1.3 Government and tax instruments

There is a benevolent government whose objective it is to maximize social welfare

\[ W \equiv f_M \psi_M V_M + f_R \psi_R V_R + f_C \psi_C V_C, \]  

(6)

where \( \psi_i \) is the Pareto weight attached to workers of type \( i \), where the weights satisfy \( \sum f_i \psi_i = 1 \), and \( V_i \equiv U(c_i, \ell_i) \) are indirect utilities. While the government is aware of the structure of the economy, it cannot observe an individual’s occupation. This assumption is satisfied by real world tax systems which also do not condition taxes on occupation, for example, because enforcement may be difficult. However, the government can observe individual income and consumption, as well as the value of robots purchased by the final goods producer. Accordingly, I assume that the government has access to two tax instruments: a non-linear income tax, and a tax on the value of robots.

Denote by \( y_i \equiv w_i \ell_i \) gross labor income earned by an individual of type \( i \). The government levies a non-linear income tax \( T(y) \) on gross labor income \( y \). Taking the wage and income tax schedule as given, a worker of type \( i \) then maximizes utility (1) by choosing consumption and labor supply subject to a budget constraint:

\[ \max_{c_i, \ell_i} U(c_i, \ell_i) \text{ s.t. } c_i \leq w_i \ell_i - T(w_i \ell_i). \]  

(7)

The value of robots purchased by the final goods producer is given by \( qB \), on which the government may levy a proportional tax \( \tau \), to which I refer as robot tax. The user price of robots is then \( p = (1 + \tau) q \). While throughout this paper I refer to \( \tau \) as a tax on robots, I highlight that \( \tau \) may be negative, and may thus be a subsidy. The government faces a budget constraint

\[ f_M T(y_M) + f_R T(y_R) + f_C T(y_C) + \tau q B = 0, \]  

(8)

stating that by raising tax revenue with the income tax and taxes on robots and other capital, it must break even. Introducing an exogenous revenue requirement does not change the analysis.

3.2 Optimal policy

The government chooses tax instruments \( T(\cdot) \) and \( \tau \) such as to maximize social welfare (6) subject to budget constraint (8). To characterize optimal taxes, I follow the conventional approach of first solving for the optimal allocation from a mechanism design problem. Afterward, prices and optimal taxes that decentralize the allocation are determined. In a direct mechanism, workers announce their type \( i \), and then get assigned consumption \( c_i \) and labor supply \( \ell_i \). Here, I consider the equivalent problem in which instead of consumption, the planner allocates indirect utilities \( V_i \) and define \( c(V_i, \ell_i) \) as the inverse of \( U(c_i, \ell_i) \) with respect to its first argument.

---

*I focus on a linear tax on robots, since with a non-linear tax and constant returns to scale there would be incentives for firms to break up into parts until each part faces the same minimum tax burden. With linear taxes, such incentives are absent.
The allocation must induce workers to truthfully report their type and thus needs to be incentive compatible. Since there is no heterogeneity of types within occupations, the only way in which workers can imitate one another is by mimicking incomes of workers in other occupations. I assume that the primitives of the model are such that \( w_C > w_R > w_M \) is satisfied. Moreover, I limit attention to those cases in which only the downward adjacent incentive constraints may be binding, while all other incentive constraints are slack. This case is the relevant one for gaining intuition which carries over to the continuous-type model. To induce a cognitive worker to truthfully report his type, the following must hold

\[
V_C \geq U \left( c(V_R, \ell_R), \ell_R \frac{w_R(L, B)}{w_C(L, B)} \right),
\]

where \( \ell_R \frac{w_M}{w_C} \) is the amount of labor which a cognitive worker needs to supply to mimic the income of a routine worker. Similarly, a routine worker has to be prevented from mimicking the income of manual workers, and thus

\[
V_R \geq U \left( c(V_M, \ell_M), \ell_M \frac{w_M(L, B)}{w_R(L, B)} \right).
\]

### 3.2.1 Separation into inner and outer problem

I follow Rothschild and Scheuer (2013, 2014) and separate the mechanism design problem into an inner problem and an outer problem. In the inner problem, the planner takes the tuple of inputs \((L, B)\) as given and maximizes welfare \( W(L, B) \) over \( \{V_i, \ell_i\}_{i \in \mathcal{I}} \) subject to constraints (specified below). In the outer problem, the planner chooses the vector \( L = (L_M, L_R, L_C) \) and robots \( B \) such that \( W(L, B) \) is maximized. The mechanism design problem can thus be written as

\[
\max_{L, B} W(L, B) \equiv \max_{\{V_i, \ell_i\}_{i \in \mathcal{I}}} f_M \psi_M V_M + f_R \psi_R V_R + f_C \psi_C V_C
\]

subject to the incentive constraints (9) and (10), the consistency conditions

\[
f_i \ell_i - L_i = 0 \ \forall i \in \mathcal{I},
\]

and the resource constraint

\[
\sum_{i \in \mathcal{I}} f_i \ell_i Y_i (L, B) + Y_B (L, B) B - \sum_{i \in \mathcal{I}} f_i c_i - qB = 0.
\]

The consistency conditions (12) restate the definition of aggregate labor supplies. Since the inner and outer problem separate optimization over individual labor supplies and aggregate labor supplies, including the consistency conditions ensures that the labor market clears. The first two terms in the resource constraint (13) sum to total output \( Y(L, B) \). The final term captures that \( x = qB \) units of the final good have to be used to produce \( B \) robots.

### 3.2.2 Optimal robot tax

I first characterize the optimal robot tax by using that in the outer problem at the optimum \( \partial W(L, B) / \partial B = 0 \), hence a change in robots may not lead to a change in welfare.
Proposition 1. The optimal tax on robots is characterized by

\[ \tau q B = \varepsilon_{w_C/w_R, B} I_{CR} - \varepsilon_{w_M/w_R, B} I_{RM} \]  

with elasticities of relative wages with respect to the number of robots defined as

\[ \varepsilon_{w_C/w_R, B} = \frac{\partial (w_C/w_R)}{\partial B} \frac{B}{w_C/w_R} > 0, \]  

\[ \varepsilon_{w_M/w_R, B} = \frac{\partial (w_M/w_R)}{\partial B} \frac{B}{w_M/w_R} > 0, \]  

and incentive effects

\[ I_{CR} = f_C (1 - \psi_C) \left( \ell_R \frac{w_R}{w_C} \right)^{1+\frac{1}{\varepsilon}} \]  

\[ I_{RM} = f_M (\psi_M - 1) \left( \ell_M \frac{w_M}{w_R} \right)^{1+\frac{1}{\varepsilon}}. \]

Proof. See Appendix A.1. \qed

The left-hand side of (14), \( \tau q B \), is the tax revenue raised with the robot tax. Ceteris paribus, the robot tax is thus larger in magnitude, the smaller the cost of producing robots, \( q \), and the lower the number of robots, \( B \). At the optimum, robot tax revenue is equal to the difference in incentive effects \( I_{CR} \) and \( I_{MR} \), weighted by the respective elasticity terms, \( \varepsilon_{w_C/w_R, B} \) and \( \varepsilon_{w_M/w_R, B} \). The elasticity terms capture the percentage increase in wages of non-routine workers relative to the wage of routine workers due to a one-percent increase in the number of robots. By Assumption 1, an increase in robots raises the wage of non-routine workers relative to routine workers. As a consequence, \( \varepsilon_{w_C/w_R, B} > 0 \) and \( \varepsilon_{w_M/w_R, B} > 0 \). The incentive effects \( I_{CR} \) and \( I_{RM} \) capture how incentive constraints (9) and (10) are affected by a marginal increase in robots, and how this, in turn, affects welfare.

I first focus on \( I_{CR} \). Raising the number of robots increases \( w_C/w_R \), and since \( w_C > w_R \), wage inequality at the top of the wage distribution rises. Regular welfare weights decrease with income, leading to \( \psi_C < 1 \). The government thus attaches a lower-than-average weight to cognitive non-routine workers. In this case, it is desirable to redistribute income from cognitive non-routine workers to workers who earn less. The increase in wage inequality at the top then tightens the incentive constraint (9): cognitive non-routine workers now need to put in less labor than before to imitate the income of a routine worker. This tightening of (9) corresponds to increased income-tax distortions of labor supply, which limits redistribution and lowers welfare.

The robot tax has the opposite effect than an increase in \( B \). By increasing the user price of robots \( p \), the equilibrium number of robots falls. As a consequence, \( w_C/w_R \) drops, which corresponds to a reduction in wage inequality at the top of the wage distribution; and to a relaxation of incentive constraint (9). Relaxing (9) is welfare improving: it becomes less distortionary to use the income tax to redistribute income from cognitive non-routine workers to other workers. Ceteris paribus, a larger incentive effect \( I_{CR} \) calls for a higher tax on robots. If labor supply is more elastic (higher \( \varepsilon \)) income taxation is more distortionary. As a result, \( I_{CR} \) is larger, and so is the optimal robot tax. The weighting of incentive effect \( I_{CR} \) by elasticity \( \varepsilon_{w_C/w_R, B} \) captures how effective taxing robots is in reducing the wage gap \( w_C/w_R \).

I now turn to the second term on the right-hand side of (14). The incentive effect \( I_{RM} \) captures how reducing the wage gap between \( w_R \) and \( w_M \) affects welfare via the incentive constraint (10). With regular welfare weights we have \( \psi_M > 1 \), hence the government values

\[ \text{[11]There are only three levels of wages, } w_M, w_R, w_C. \]
redistributing income to manual non-routine workers. In this case, lowering the gap between \( w_R \) and \( w_M \) relaxes (10), and makes it less distortionary to redistribute income with the income tax, captured by \( I_{RM} > 0 \). The weighting with elasticity \( \varepsilon_{w_M/w_R,B} > 0 \) captures how effective the robot tax is in changing the wage gap between \( w_R \) and \( w_M \). The minus sign is crucial: a tax on robots increases the wage of routine workers relative to manual non-routine workers. As inequality at the bottom of the wage distribution increases, the incentive constraint (10) tightens. Redistributing to manual non-routine workers with the income tax becomes more distortionary, thereby lowering welfare. This effect, ceteris paribus, calls for a lower tax on robots.

To summarize: it is welfare-maximizing to distort the price of robots to make income redistribution less distortionary – and to thereby violate production efficiency (Diamond and Mirrlees, 1971). A tax on robots decreases wage inequality between cognitive and routine workers, thereby reducing income-tax distortions of labor supply and increasing welfare. At the same time, a tax on robots raises the wage gap between routine and manual workers, which worsens income-tax distortions of labor supply and lowers welfare. Due to these opposing forces, the sign of the robot tax is *ambiguous*. If the first effect dominates, robots should be taxed, whereas if the second effect is more important, robots should be subsidized.

Ceteris paribus, several factors make it more likely for the optimal robot tax to be positive: if robots increase wage inequality at the top of the distribution more than they reduce inequality at the bottom; if the share of cognitive non-routine workers is large, whereas the share of manual non-routine workers is small; if the wage gap between cognitive and routine workers is small, whereas the wage gap between routine and manual workers is large; finally, if the government attaches relatively little weight \( \psi_C \) and \( \psi_M \) to cognitive workers and manual workers, respectively. The final point can be restated as the government attaching relatively more weight \( \psi_R \) to routine workers. This is intuitive: as routine workers gain relative to non-routine workers when robots are taxed, putting relatively more weight on them calls for a larger tax on robots.

Only in special cases is the tax on robots zero. First, it is zero if the government can condition income taxes on occupation, which restores production efficiency (Diamond and Mirrlees, 1971). Moreover, since in the simple model worker types and occupations coincide, occupation-specific income taxes correspond to individualized lump-sum taxes, leading to the first-best outcome. Similarly, the optimal robot tax is zero if income taxation does not distort labor supply, corresponding to \( \varepsilon \to 0 \). The robot tax is also zero, if the effect of reducing labor-supply distortions at the top of the wage distribution exactly cancels against the effect of raising labor-supply distortions at the bottom. Finally, if in contrast to what I have assumed so far, robots are equally complementary to labor in all occupations, the optimal robot tax is zero, since in this case \( \varepsilon_{w_C/w_R} = \varepsilon_{w_M/w_M} = 0 \).

The ambiguous sign of the robot tax is in contrast to Guerreiro et al. (2017) who argue that the robot tax should be positive. This is due to Guerreiro et al. considering only two groups of workers: routine and non-routine. In their model, taxing robots unambiguously relaxes the single binding incentive constraint. My result highlights that aggregating workers into just two groups can be misleading as it masks heterogeneous effects of robots on wages along the income distribution. The empirical literature (see e.g. Autor and Dorn, 2013) finds that routine workers are not found at the very bottom of the income distribution. Instead, those who earn least often perform non-routine manual work which is hard to automate. In this case, taxing robots will widen inequality at the bottom of the wage distribution, thereby worsening income-tax distortions of labor supply. If these effects are taken into account, the sign of the robot tax becomes ambiguous.

\[12\] Along the same lines, Slavík and Yazici (2014) show that structures capital which is equally complementary to labor in all occupations should not be taxed.
3.2.3 Optimal income taxes

I now use the inner problem, taking L and B as given, to characterize optimal marginal income taxes.

**Proposition 2.** Let $\mu$ denote the multiplier on the resource constraint (13). Let $\mu_{\eta CR}$ be the multiplier on incentive constraint (9) and $\mu_{\eta RM}$ the multiplier on (10). Define $\mu_{\xi}$ as the multiplier on the consistency condition for occupation $i$ in (12). The optimal marginal income tax rates satisfy

$$
\frac{T'_M + \frac{\xi_M}{Y_M}}{1 - T'_M} = (\psi_M - 1) \left(1 - \left(\frac{w_M}{w_R}\right)^{1+\frac{1}{\varepsilon}}\right)
$$

(19)

$$
\frac{T'_R + \frac{\xi_R}{Y_R}}{1 - T'_R} = \frac{f_C}{f_R} (1 - \psi_C) \left(1 - \left(\frac{w_R}{w_C}\right)^{1+\frac{1}{\varepsilon}}\right)
$$

(20)

$$
T'_C = \frac{\xi_C}{Y_C},
$$

(21)

with

$$
\xi_i = \tilde{\varepsilon}_{wR/wC}I_{CR} + \tilde{\varepsilon}_{wM/wR}I_{RM},
$$

(22)

where the semi-elasticities of relative wages with respect to $L_i$ are defined as

$$
\tilde{\varepsilon}_{wR/wC}I_{CR} \equiv \frac{\partial}{\partial L_i} \left(\frac{w_R}{w_C}\right) \frac{w_C}{w_R}
$$

(23)

and

$$
\tilde{\varepsilon}_{wM/wR}I_{RM} \equiv \frac{\partial}{\partial L_i} \left(\frac{w_M}{w_R}\right) \frac{w_R}{w_M}.
$$

(24)

**Proof.** See Appendix A.2

First, note that each expression for optimal marginal income tax rates features a correction for general-equilibrium effects, $\xi_i/Y_i$. Suppose for the moment that general-equilibrium effects are absent. We then have $\xi_i = 0 \forall i \in I$. Moreover, assume that welfare weights satisfy $\psi_M > 1$ and $\psi_C < 1$, as will be the case with a welfarist government which attaches higher weights to individuals who earn lower incomes. As a consequence, marginal tax rates $T'_M$ and $T'_R$ are positive.\(^{13}\) This is in line with the function of marginal income tax rates: the role of the marginal tax rate at income $y$ is to redistribute income from individuals who earn more than $y$ to individuals earning equal to, or less than, $y$. Consider (19): the social marginal value of distributing income from $f_R$ routine workers and $f_C$ cognitive workers to a mass of $f_M$ manual workers is $(\psi_M - 1) = \tilde{\varepsilon} M \epsilon_j [f_R (1 - \psi_R) + f_C (1 - \psi_C)]$. Similarly, in (20), the term $f_R^{-1} f_C (1 - \psi_C)$ captures the social marginal value of redistributing income from $f_C$ cognitive workers to a mass of $f_R$ routine workers. However, marginal tax rates distort labor supply, which is captured by the terms $1 - (w_M/w_R)^{1+1/\varepsilon}$ and $1 - (w_R/w_M)^{1+1/\varepsilon}$. With $\varepsilon > 0$, both terms are smaller than 1 and thus scale down marginal tax rates, whereas in the absence of labor supply responses $\varepsilon \to 0$, and both terms tend to 1. Finally, without general-equilibrium effects $T'_C = 0$. This is the famous result of “no distortion at the top” (Sadka, 1976; Seade, 1977). Since there are no individuals who earn more than cognitive workers, setting a positive marginal tax rate has no distributional benefits but would distort labor supply. It is thus optimal to set a marginal tax rate of zero.

Now consider the case with general-equilibrium effects. Under special conditions, it is possible to sign the multiplier terms $\xi_M$ and $\xi_C$.

\(^{13}\)Note that there are three discrete income levels in the economy. The income tax function $T$ is thus not differentiable. Marginal tax rates $T'$ are defined as $T_i \equiv 1 + \frac{\partial L_i}{\partial L_i}(\psi_i) \frac{1}{w_i}$.
Corollary 1. With general-equilibrium effects, the multiplier terms $\xi_M$ and $\xi_C$ can be signed as follows:

- if $\frac{\partial}{\partial L_M} \left( \frac{w_B}{w_C} \right) < 0$, $\xi_M < 0$,
- if $\frac{\partial}{\partial L_C} \left( \frac{w_M}{w_R} \right) > 0$, $\xi_C > 0$.

Proof. See Appendix A.2.4

Suppose that $\tilde{\xi}_{wR/wC,L_M} < 0$ and thus $\xi_M < 0$. In this case $T_M'$ is larger than in the absence of general-equilibrium effects. The intuition has again to do with the relaxation of incentive constraints. A higher marginal tax rate $T_M'$ discourages labor supply of manual workers, which increases their wage relative to routine workers, thereby relaxing incentive constraint (9). Moreover, since by assumption $\tilde{\xi}_{wR/wC,L_M} < 0$, a reduction in the supply of manual labor also increases the wage of routine workers relative to cognitive workers, which relaxes incentive constraint (9). A similar reasoning applies to the case in which $\tilde{\xi}_{wM/wR,L_C} > 0$ and thus $\xi_C > 0$. Now $T_C'$ becomes negative, as in Stiglitz (1982), which encourages labor supply of cognitive workers. As a result, their wage drops relative to routine workers, whereas, by assumption, the wage of manual workers increases relative to routine workers. Again, this overall wage compression relaxes incentive constraints, and is therefore welfare improving. Whether the marginal income tax for routine workers is scaled up or down in the presence of general-equilibrium effects is ambiguous. Consider an increase in $T_R'$: As a result, labor supply of routine workers falls, raising their wage relative to manual and cognitive workers. This change has opposing effects on incentive constraints: it relaxes (9) but tightens (10). Whether an increase in $T_R'$ is desirable thus depends on which of the two effects is more relevant for welfare.

3.2.4 Effect of robot tax on marginal income taxes

How are marginal income taxes affected by the presence of the robot tax? To answer this question, I first derive expressions for the case in which taxing robots is not possible. To do so, I impose the additional constraint $Y_B(L, B) = q$, which corresponds to $\tau = 0$. The expressions in (19), (20) and (21) are not affected by the absence of the robot tax. However, the multipliers on the consistency conditions are now different.

Corollary 2. In the absence of the robot tax, let $\mu \kappa$ denote the multiplier on the additional constraint $Y_B(L, B) − q = 0$. Let $\mu_i \xi_i$ be the multiplier on the consistency condition for occupation $i$, $\mu_{iCR}$ the multiplier on incentive constraint (9) and $\mu_{iRM}$ the multiplier on incentive constraint (10). The following condition holds for $\xi_i$:

$$\xi_i = \tilde{\xi}_{wR/wC,L_i} I_{CR} + \tilde{\xi}_{wM/wR,L_i} I_{RM} + \kappa \frac{\partial Y_B(L, B)}{\partial L_i} \forall i, \quad (25)$$

with

$$\kappa \frac{\partial Y_B(L, B)}{\partial B} B = \tilde{\xi}_{wC/wR,B} I_{CR} - \tilde{\xi}_{wM/wR,B} I_{MR}. \quad (26)$$

Without the robot tax, $\xi_i$ is thus adjusted by $\kappa \partial Y_B(L, B)/\partial L_i$. Note that the right-hand-side of (26) is the same as in the expression for the optimal robot tax (14). Since $\partial Y_B(L, B)/\partial B < 0$, $\kappa$ thus has the opposite sign of the optimal robot tax which would result if we were not to rule out robot taxation. Suppose that the optimal robot tax would be positive, and thus $\kappa < 0$. I first focus on the unambiguous cases $i \in \{M, C\}$. We then have $\partial Y_B(L, B)/\partial L_i > 0$ and hence $\xi_i$ is lower in the absence of the robot tax. It thus follows that with the robot tax, both $T_M'$ and $T_C'$ are lower. Intuitively, a tax on robots lowers the wages of manual and cognitive workers, which induces them to reduce their labor supply. Lower marginal
income tax rates encourage labor supply, which partly offset the reduction. Moreover, in the 
case of cognitive workers, the drop in wages which results from increased labor supply further 
compresses the wage distribution. Now consider routine workers. The sign of \( \frac{\partial Y_B (L, B)}{\partial L_R} \) is ambiguous, and as a result it is not clear in which direction the marginal income tax for 
routine workers is adjusted if robots are taxed. Suppose that \( \frac{\partial Y_B (L, B)}{\partial L_R} > 0 \), which will 
be the case if the difference in elasticities of substitution between robots and routine workers 
on the one hand and robots and cognitive or manual workers on the other hand is not too large. 
\( \xi_R \) is now lower without the robot tax, and thus, \( T_R' \) will be lower if robots can be taxed.

4 Continuous types and occupational choice

Having developed intuition, I now extend the model to allow for continuous types as well as 
for occupational choice. To do so, I build on the framework by Rothschild and Scheuer (2013, 
2014).

4.1 Setup

4.1.1 Skill Heterogeneity

There is a unit mass of individuals. Each individual is characterized by three-dimensional 
skill-vector \( \theta \in \Theta \equiv \Theta_M \times \Theta_R \times \Theta_C \), with \( \Theta_i \equiv [\underline{\theta}_i, \bar{\theta}_i] \) and \( i \in I \equiv \{M, R, C\} \). As before, 
each dimension of skill determines an individual’s productivity in one of the three occupations: manual non-routine (\( M \)), routine (\( R \)), and cognitive non-routine (\( C \)). Skills are distributed 
according to a continuous cumulative distribution function \( F : \Theta \rightarrow [0,1] \) with corresponding 
density \( f \).

4.1.2 Technology

The final good is produced using aggregate effective labor in the three occupations, \( L \equiv (L_M, L_R, L_C) \), and robots \( B \) according to a constant returns to scale production function 
\( Y (L, B) \). As before, robots are better substitutes for routine labor than for non-routine la-
bor, and Assumption 1 holds. Aggregate effective labor is now defined as

\[
L_M \equiv \int_M \theta_M \ell (\theta) dF (\theta), \quad L_R \equiv \int_R \theta_R \ell (\theta) dF (\theta), \quad L_C \equiv \int_C \theta_C \ell (\theta) dF (\theta),
\]

(27)

where \( M, R \) and \( C \) are the sets of individuals \( \theta \) working in occupations \( M, R \) and \( C \), respectively. Robots are produced linearly with the final good, as described in Section 3.1.2.

4.1.3 Preferences and occupational choice

Individuals derive utility from consumption \( c \) and disutility from supplying labor \( \ell \) according to the strictly concave utility function \( U(c, \ell) \) with \( U_c > 0, U_\ell < 0 \). Let \( y \) denote an individual’s gross income and \( w \) the wage. I assume that \( U \) satisfies the standard Spence-Mirrlees single-crossing property (Mirrlees, 1971; Ebert, 1992; Hellwig, 2004), that is, the marginal rate of substitution between income and consumption, \( -U_\ell (c, \frac{y}{w}) / (wU_c (c, \frac{y}{w})) \), decreases in \( w \). Moreover, I assume that the monotonicity condition is satisfied, that is, gross-income needs to increase in \( w \).\(^{14}\) Individuals choose their occupation according to a Roy (1951) model, such that their wage is maximized and, in equilibrium, given by

\[
w_{L,B}(\theta) = \max \{ Y_M (L, B) \theta_M, Y_R (L, B) \theta_R, Y_C (L, B) \theta_C \},
\]

(28)

\(^{14}\)With non-linear taxes, first-order conditions are necessary, but generally not sufficient for utility maximization. The single-crossing and monotonicity conditions ensure that second-order conditions for utility maximization hold.
the overall density is wages, $\Psi$, there is a unique mapping from types to wages, Pareto weights can be written as function of $\theta$ types, which I now denote by $\Psi(\theta)$. As in Section 3, I separate the planner problem into an inner problem over labor supply and indirect utilities, subject to incentive and resource constraints. Let the wage density for occupation $L$ be $\omega_L(w) = \frac{w}{Y_L(L,B)}$. Moreover, once $L, B$ is fixed, all endogenous variables which depend on an individual’s type can be written in terms of wages. This feature allows for a useful separation of the planner problem.

4.2.2 Separation into inner and outer problem

As in Section 3, I separate the planner problem into an inner problem which maximizes welfare over labor supply and indirect utilities for given $L$ and $B$, and an outer problem which maximizes welfare over $L$ and $B$.

### Inner problem
Denote by $V(\theta)$ the indirect utility of type $\theta$. Social welfare is defined as an integral over weighted indirect utilities $\int_{\theta} V(\theta) d\Psi(\theta)$, which given $L, B$ can be written as $\int_{\theta_L} V(w) d\Psi_{L,B}(w)$. As is common, I maximize social welfare by directly choosing an allocation of indirect utilities and labor supplies, subject to incentive and resource constraints.

### Separation into inner and outer problem

I follow Rothschild and Scheuer (2013) to reduce the dimensionality of the problem: Given factor inputs $L$ and $B$, wage rates and sectoral choice are determined, and the three-dimensional heterogeneity in skill can be reduced to one-dimensional heterogeneity in wages. The distribution of skills $F(\theta)$ corresponds to the wage distribution

$$F_{L,B}(w) = F\left(\frac{w}{Y_M(L,B)}, \frac{w}{Y_R(L,B)}, \frac{w}{Y_C(L,B)}\right)$$

with occupational wage densities

$$f^M_{L,B}(w) = \frac{1}{Y_M(L,B)} \int_{\theta_C}^{w/Y_M(L,B)} f\left(\frac{w}{Y_M(L,B)}, \theta_M, \theta_C\right) d\theta_M d\theta_C,$$

$$f^R_{L,B}(w) = \frac{1}{Y_R(L,B)} \int_{\theta_C}^{w/Y_R(L,B)} f\left(\frac{w}{Y_R(L,B)}, \theta_R, \theta_C\right) d\theta_R d\theta_C,$$

$$f^C_{L,B}(w) = \frac{1}{Y_C(L,B)} \int_{\theta_R}^{w/Y_C(L,B)} f\left(\frac{w}{Y_C(L,B)}, \theta_R, \theta_C\right) d\theta_R d\theta_C.$$
In addition, the allocation needs to be consistent with \( L \) and \( B \) to make sure that the market for robots and labor clears. Factor market clearing is ensured by consistency conditions. Apart from these consistency conditions, the problem is a standard Mirrlees (1971) problem. I define the inner problem as

\[
W(L, B) \equiv \max_{V(w), \ell(w)} \int_{\mathcal{W}_{L,B}} V(w) \, d\Psi_{L,B}(w)
\]

subject to

\[
V'(w) + U_{\ell}(c(V(w), \ell(w))) \cdot \frac{\ell(w)}{w} = 0 \quad \forall w \in [w_{L,B}, \bar{w}_{L,B}]
\]

\[
\frac{1}{Y_M(L, B)} \int_{\mathcal{W}_{L,B}} w\ell(w) f_{L,B}^M(w) \, dw - L_M = 0
\]

\[
\frac{1}{Y_R(L, B)} \int_{\mathcal{W}_{L,B}} w\ell(w) f_{L,B}^R(w) \, dw - L_R = 0
\]

\[
\frac{1}{Y_C(L, B)} \int_{\mathcal{W}_{L,B}} w\ell(w) f_{L,B}^C(w) \, dw - L_C = 0
\]

\[
\int_{\mathcal{W}_{L,B}} (w\ell(w) - c(V(w), \ell(w))) \, f_{L,B}(w) \, dw + Y_B(L, B) B - qB = 0.
\]

Here, (34) is the set of incentive constraints, (35), (36), (37) are the consistency conditions for occupations \( M \), \( R \) and \( C \), respectively. Equation (38) is the resource constraint and the continuous-type equivalent of (13).

**Outer problem.** In the outer problem, the planner chooses inputs \( L \) and \( B \) such that welfare is maximized, that is, he solves \( \max_{L,B} W(L, B) \). It is useful that \( W(L, B) \) corresponds to the value of the Lagrangian of the inner problem, evaluated at optimal indirect utilities and labor supplies.

### 4.2.3 Optimal robot tax

I obtain a condition for the optimal robot tax from the outer problem by differentiating the maximized Lagrangian with respect to robots, \( B \). As in the simple model, the optimal allocation can be implemented using a linear tax on the value of robots. To see this, note that since firms maximize profits, they equate the marginal return to robots with the price of robots. Under laissez-faire, we thus have

\[
Y_B(L, B) = q.
\]

However, given \( L \), the planner might want to distort the choice of robots such that (39) is no longer satisfied. By setting a linear tax \( \tau \) on the value of robots, profit maximization of the firm leads to

\[
Y_B(L, B) = (1 + \tau) q.
\]

Since \( Y_B(L, B) \) is strictly monotone in \( B \), for each \( B \) there exists a unique robot tax \( \tau \) such that (40) holds. For given optimal \( L \), the optimal robot tax \( \tau \) thus uniquely implements the optimal \( B \). Using (40), I characterize the optimal robot tax as follows.

**Proposition 3.** Let \( \mu \) denote the multiplier on the resource constraint and \( \mu \xi_i \) the multiplier on the consistency condition for occupation \( i \in \{M, R, C\} \). Let \( \mu(w) \) denote the multiplier on
the incentive constraint. Denote by \( q^i_E \) the income share in occupation \( i \). The optimal tax on robots, \( \tau^* \), is characterized by

\[
\tau q_B = \varepsilon_{Y_C/Y_R,B}(L, B) \left( I_C(L, B) + \sum_{i \in I} \xi_i (C_{C_i}(L, B) + S_{C_i}(L, B)) \right) + \varepsilon_{Y_M/Y_R,B}(L, B) \left( I_M(L, B) + \sum_{i \in I} \xi_i (C_{M_i}(L, B) + S_{M_i}(L, B)) \right),
\]

(41)

where for \( i \in \{ M, R, C \} \)

\[
I_i(L, B) \equiv \int_{w_{L,B}}^{w_{L,B}} \eta(w) u'(w) w \frac{d}{dw} \left( \frac{f_{L,B}(w)}{f_{L,B}(w)} \right) dw,
\]

(42)

and for \( i, j \in \{ M, R, C \} \)

\[
C_{ij}(L, B) \equiv \frac{1}{Y_j(L, B)} \int_{w_{L,B}}^{w_{L,B}} w^2 \ell'(w) \text{Cov} \left( q^i_{L,B}(\theta), q^j_{L,B}(\theta) | w \right) f_{L,B}(w) dw,
\]

(43)

with

\[
q^i_{L,B}(\theta) = \begin{cases} 
1, & \text{if } \theta \text{ works in } i \\
0, & \text{otherwise,}
\end{cases}
\]

and \( S_{ij} \) as defined in the Appendix. The elasticities of equilibrium wage rates with respect to the amount of robots are defined as

\[
\varepsilon_{Y_M/Y_R,B}(L, B) \equiv \frac{\partial (Y_M(L, B) / Y_R(L, B))}{\partial B} \frac{B}{Y_M(L, B) / Y_R(L, B)} > 0,
\]

(44)

and

\[
\varepsilon_{Y_C/Y_R,B}(L, B) \equiv \frac{\partial (Y_C(L, B) / Y_R(L, B))}{\partial B} \frac{B}{Y_C(L, B) / Y_R(L, B)} > 0.
\]

(45)

Proof. See Appendix B

The expression in (41) characterizes the optimal tax revenue raised with the robot tax. First, note the similarity between (41) and the corresponding expression in the simplified model, (14). In both cases, the effect of robots on relative wage rates plays a crucial role.\footnote{In the simplified model elasticities of relative wages coincide with elasticities of relative wage rates. Due to heterogeneity of wages within occupations, this is no longer the case here.} Due to Assumption 1, an increase in the number of robots leads to higher wage rates in non-routine occupations relative to routine occupations. As a consequence, the elasticities of relative wage rates \( \varepsilon_{Y_M/Y_R,B}(L, B) \) and \( \varepsilon_{Y_C/Y_R,B}(L, B) \) are both positive.

As in the simple model, these elasticities multiply the incentive effects \( I_i \). In addition, they multiply terms which emerge due to heterogeneous types and occupational choice. Following Rothschild and Scheuer (2013), I refer to these as effort reallocation effects \( C_{ij} \) and occupational shift effects \( S_{ij} \), where \( i, j \in \{ M, R, C \} \). Incentive effects, effort reallocation effects and occupational shift effects ultimately affect welfare for the same reason: a change in relative wage rates leads to a change in the wage distribution, which affects incentive constraints – and thus – income-tax distortions of labor supply. If income-tax distortions are reduced, more income can be redistributed overall, which raises welfare.

It is instructive to think about incentive effects as capturing the first-round welfare impact of a tax on robots on relative wage rates. In response to changed relative wage rates, individuals
adjust their behavior, which then has second-round effects on relative wage rates and on welfare. These second-round effects are captured by the effort reallocation and occupational shift effects. They originate from the effect of robots on the consistency conditions. Intuitively, reallocation and occupational shift effects reduce the effectiveness of the robot tax, by counteracting initial wage compression.

**Incentive effects.** The incentive effects capture how changes in relative wage rates affect tax-distortions on labor supply, and thus welfare. If \( I_i(\mathbf{L}, \mathbf{B}) > 0 \), an increase in the robot tax leads to welfare-improving wage compression. Ceteris paribus, larger incentive effects thus call for a higher tax on robots. To determine the sign of the incentive effect, suppose that the incentive constraint (34) is downward-binding, and thus \( \eta(w) \geq 0 \). Since indirect utilities \( V(w) \) increases in \( w \), the sign of \( I_i(\mathbf{L}, \mathbf{B}) \) is determined by \( \frac{d}{dw} \left( \frac{f_{L,B}(w)}{f_{L,B}(w)} \right) \). This term captures how the share of individuals earning wage \( w \) in occupation \( i \) changes with a marginal increase in \( w \).\(^{18}\)

Consider \( I_C \). Since workers in cognitive occupations are concentrated at high wages, the term \( \frac{d}{dw} \left( \frac{f_{C,B}^C(w)}{f_{L,B}(w)} \right) \) is positive at most \( w \). As a consequence, we find \( I_C > 0 \). By reducing \( Y_C / Y_R \), a tax on robots thus compresses wages at the top of the wage distribution, which increases welfare. The intuition is similar as in the stylized model. Wage compression at the top of the distribution makes it more costly for cognitive workers to imitate types who earn marginally lower incomes in routine occupations – which locally relaxes incentive constraints. In other words, income-tax distortions of labor supply are locally alleviated. This allows for more redistribution overall, which raises welfare.

Next, consider \( I_M \) and suppose that manual non-routine workers are concentrated at low wages, as observed empirically. The term \( \frac{d}{dw} \left( \frac{f_{M,B}^M(w)}{f_{L,B}(w)} \right) \) is then negative at most \( w \), leading to \( I_M < 0 \). The negative sign captures that a tax on robots lowers welfare by locally tightening incentive constraints at the bottom of the wage distribution. As in the stylized model, a tax on robots thus has opposing effects on incentive constraints – and thus labor-supply distortions – at the top and at the bottom of the wage distribution. As a consequence, the sign of the robot tax is again ambiguous.

**Effects on the consistency conditions.** Both, effort reallocation and occupational shift effects capture changes in aggregate labor supplies, which affect welfare via the consistency conditions. The welfare impact of a marginal increase in \( L_i \) via the consistency condition is \( \mu \xi_i \), with \( \mu > 0 \). The multiplier \( \mu \xi_i \) thereby captures how the change in \( L_i \) affects welfare by changing relative wage rates, and as a consequence incentive constraints. Different from the simple model, it is not anymore possible to sign \( \xi_i \) analytically.

**Effort reallocation effects.** The effort reallocation effect \( C_{ij} \) captures the welfare-relevant impact of a marginal increase in \( Y_i \) on aggregate labor supply \( L_j \) which arises due to individuals

\(^{17}\)In the model, equilibrium is determined simultaneously, hence there are no actual first and second rounds.

\(^{18}\)Note that if \( I_i \geq 0 \) for some occupation \( i \), it has to hold that there is at least one other occupation \( j \neq i \), for which \( I_j \leq 0 \). To see this, write

\[
\begin{align*}
I_i(\mathbf{L}, \mathbf{B}) &= \int_{w_{L,B}}^{w_{L,B}} \eta(w) \left( \frac{U_c(w)}{U_c(w)} \right) V'(w) w \frac{d}{dw} \left( \frac{f_{L,B}(w)}{f_{L,B}(w)} \right) dw \\
&= \int_{w_{L,B}}^{w_{L,B}} \eta(w) \left( \frac{U_c(w)}{U_c(w)} \right) V'(w) w \frac{d}{dw} \left( \frac{f_{L,B}(w)}{f_{L,B}(w)} - \sum_{j \neq i} f_{L,B}(w) \right) dw \\
&= - \sum_{j \neq i} \int_{w_{L,B}}^{w_{L,B}} \eta(w) \left( \frac{U_c(w)}{U_c(w)} \right) V'(w) w \frac{d}{dw} \left( \frac{f_{L,B}(w)}{f_{L,B}(w)} \right) dw.
\end{align*}
\]
adjusting their labor supply within occupations, while keeping occupational choice fixed.

Recall that we derive the expression for the optimal robot tax from the outer problem, taking as given \( \ell(w) \) which is chosen optimally in the inner problem. Still, by affecting relative wage rates, a change in the number of robots has an impact on labor supplies within occupations since individuals move along the schedule \( \ell(w) \), leading to a change in the wage density \( f(w) \) at \( w \). Instead of deriving changes in \( L_j \) by keeping \( \ell(w) \) fixed, and adjusting the densities, I follow Rothschild and Scheuer (2014) and construct a variation of the \( \ell(w) \) schedule which, at each \( w \), neutralizes average changes in \( \ell(w) \) across occupations. As a result, the wage density \( f(w) \) is unaffected. Moreover, at the margin, the schedule variation has no effect on welfare.

Consider for example \( C_{CM} \). The term captures how an increase in \( Y_C \) – ceteris paribus – affects aggregate labor supply \( L_M \) due to effort reallocation. At each \( w \), individuals increase their labor supply in occupation \( C \), whereas labor supply remains unchanged in occupations \( M \) and \( R \). As a result, at each \( w \), labor supply increases more than average in \( C \) and less than average in \( M \). Intuitively, this negative correlation of changes in labor supplies in \( C \) and \( M \) at \( w \) gives rise to the covariance term in \( C_{CM} \). After neutralizing average changes, \( L_M \) decreases, which is captured by \( C_{CM} < 0 \).

**Occupational shift effects.** The set of occupational shift effects \( S_{ij}(L,B) \) capture the welfare-relevant impact of a marginal increase in \( Y_i \) on aggregate labor supply \( L_j \) due to individuals switching between occupations \( i \) and \( j \), while keeping the labor supply schedule \( \ell(w) \) fixed and ruling out shifts along \( \ell(w) \) within occupations. Since a marginal increase in \( Y_i \) lets individuals shift from occupation \( j \) to occupation \( i \) (with \( i \neq j \)), aggregate labor supply \( L_j \) is reduced. This is captured by \( S_{ij} < 0 \) for \( i \neq j \). In contrast, \( S_{ii} > 0 \) accounts for the inflow into occupation \( i \) due to an increase in \( Y_i \), leading to increase in \( L_i \).

### 4.2.4 Marginal income taxes

Turning to optimal marginal income taxes, I note that the only difference with the inner problem in Rothschild and Scheuer (2014) is the presence of the term \( q_B \) in the resource constraint. In the inner problem this term is kept fixed and does therefore not influence the optimal allocation of indirect utilities and labor. As a consequence, the characterization of optimal marginal income taxes is the same as in Rothschild and Scheuer (2014).

**Proposition 4.** Let \( \mu \) be the multiplier on the resource constraint and denote by \( \mu \xi_i \) the multiplier on the consistency condition pertaining to occupation \( i \). Let \( \varepsilon^u \) be the uncompensated labor supply elasticity and \( \varepsilon^c \) the compensated labor supply elasticity. Given \( L \) and \( B \), optimal marginal tax rates are characterized by

\[
1 - T'(w) = \left( 1 - \sum_{i \in I} \frac{\xi_i}{Y_i(L,B)} \frac{f_{L,B}^{i,j}(w)}{f_{L,B}(w)} \right) \left( 1 + \frac{\eta(w)}{\eta(w) + 1 + \varepsilon^u(w)} \right)^{-1}
\]

with

\[
\eta(w) = \int_{w_{L,B}}^{\infty} \left( 1 - \psi_{L,B}(z) U_c(z) \mu \right) \exp \left( \int_w^z \left( 1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)} \right) \frac{dy(s)}{y(s)} \right) f_{L,B}(z) \, dz.
\]

**Proof.** See Rothschild and Scheuer (2014) \( \square \)

As highlighted by Rothschild and Scheuer (2014), the formula closely resembles the optimal income tax expression in the standard Mirrlees model. The only difference is the correction

\( ^{19} \text{If, instead, one wanted to compute the total change in income due to individuals switching, one would also have to account for changes in labor supply within occupations. However, these effects have already been taken into account in the effort reallocation effects.} \)
term \(1 - \sum_{i \in I} \frac{\xi_i f_L B(w)}{Y_i(L, B)}\), which adjusts retention rates \(1 - T'(w)\) to account for general-equilibrium effects. The fact that the expression for optimal marginal income tax rates in my model is the same as in Rothschild and Scheuer (2014) does not mean that taxes on robots and other capital do not interact with optimal income taxes. As in the simple model, the presence of these taxes changes the multipliers \(\xi_i\). Since it not possible to solve for \(\xi_i\), I only study the effect of the robot tax on marginal income taxes as part of the quantitative analysis below.

5 Quantitative analysis

In this section, I study optimal taxes on robots and labor income quantitatively. To do so, I first calibrate the model to the US economy for the existing tax system. The calibrated model is then used for optimal tax analysis. The theoretical results are derived using a very general definition of robots, expressed in Assumption 1. Here, I focus on taxing a specific type of robots: industrial robots.\(^{20}\) Industrial robots are an important automation technology. Moreover, their impact on the economy has been studied empirically. Acemoglu and Restrepo (2017) analyze the impact of industrial robots on employment and wages in the US. I use their findings to guide the quantitative analysis. Acemoglu and Restrepo base their measure of robots on data from the International Federation of Robotics, which is available from 1993 onward. Their results for the impact of robots on the labor market is based on changes in wages and employment between 1993 and 2007. I therefore choose 1993 as the base year for the calibration.

The calibration proceeds in several steps. First, I obtain the skill distribution \(F(\theta)\) as well as wage rates \(Y_i\) from data on wages and occupations. Then, based on \(F(\theta)\) and \(Y_i\), and based on assumptions on individual labor supply, aggregate labor supplies \(L_i\) are computed for a given parametric tax system. After this step, five parameters of the production function remain to be calibrated. Moreover, the initial value of robots needs to be set. Using \(L_i\) as inputs, the parameters are calibrated against three internal targets: \(Y_M, Y_R,\) and \(Y_C;\) and two external targets: the elasticity of the average routine wage with respect to robots, and the price-elasticity of firm’s use of robots. The first-order condition for the firm’s demand for robots pins down the initial amount of robots. By matching wage rates \(Y_i,\) the model generates a realistic wage distribution. Moreover, by matching the external targets, the model captures the wage response to an increase in robots, as well as the response of firms to changes in robot prices – and thus to a tax on robots.

5.1 Skill-distribution

I follow Heckman and Sedlacek (1990) and Heckman and Honoré (1990) by assuming that skills \(\theta\) follow a joint log-Normal distribution. Moreover, I normalize the mean to be the zero vector. Log skills are then distributed according to

\[
\begin{bmatrix}
\ln \theta_M \\
\ln \theta_R \\
\ln \theta_C
\end{bmatrix}
\sim N\left(0, \begin{bmatrix}
\sigma_M^2 & \rho_{MR} \sigma_M \sigma_R & \rho_{MC} \sigma_M \sigma_C \\
\rho_{MR} \sigma_M \sigma_R & \sigma_R^2 & \rho_{RC} \sigma_R \sigma_C \\
\rho_{MC} \sigma_M \sigma_C & \rho_{RC} \sigma_R \sigma_C & \sigma_C^2
\end{bmatrix}\right),
\]

(48)

where I denote by \(\sigma_i\) the standard deviation of latent skill for occupation \(i\) and by \(\rho_{ij}\) the correlation between skills in occupations \(i\) and \(j\). To estimate the parameters in (48), I use that according to the model, log wages in occupation \(i\) are given by

\[
\ln w_i = \ln Y_i + \ln \theta_i,
\]

(49)

\(^{20}\) An industrial robot is defined as “an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which can be either fixed in place or mobile for use in industrial automation applications.” See https://ifr.org/img/office/Industrial_Robots_2016_Chapter_1_2.pdf
where $Y_i$ is the equilibrium wage rate in occupation $i$. Normalizing the mean of log skills to zero, I can estimate both wage rates and the parameters in (48) using data on wages and occupations.\footnote{Without the normalization of a zero mean of log skills, wage rates cannot be identified from data on wages and occupational choice alone, as log wage rates cannot be distinguished from means of the log skill-distribution. I use wage rates as internal calibration targets. They do not have an interpretation in themselves.}

According to the Roy model, an individual chooses the occupation in which he earns the highest wage. Selection thus needs to be taken into account when deriving the distribution of observed wages within occupations. For example, due to selection, the mean of observed log wages does not correspond to the mean log wage rate. Bi and Mukherjea (2010) provide the distribution of the minimum of a tri-variate Normal distribution. The log-likelihood function for estimating the parameters of (48) is provided in Appendix C.

5.1.1 Data

To estimate the skill-distribution, I use data on wages and occupational choice from the CPS Merged Outgoing Rotation Groups (MORG) as prepared by the National Bureau of Economic Research (NBER).\footnote{See http://www.nber.org/data/morg.html} The data cover the years from 1979 to 2016. I focus on the year 1993. I restrict the sample to individuals of age 16 to 64. Hourly wages are based on weekly earnings divided by usual hours worked per week, where I exclude those individuals who work less than 35 hours. Like Autor and Dorn (2013), I correct for top-coded weekly earnings by multiplying them by 1.5. All wages are converted into 2016 dollar values using the personal consumption expenditures chain-type price index.\footnote{I obtain the price index from https://fred.stlouisfed.org/series/DPCERG3A086NBEA} Occupations are categorized into three groups: manual non-routine, routine, and cognitive non-routine following Acemoglu and Autor (2011).\footnote{Acemoglu and Autor (2011) use the labels ‘abstract’ for cognitive non-routine and ‘services’ for manual non-routine. See Cortes, 2016 for the same classification and labels as used in this paper.} I provide an overview of the occupations contained in the three categories in Table 1 as well as summary statistics in Table 2. In the data, close to 60% of individuals work in routine occupations, followed by about 30% in cognitive non-routine occupations. I highlight that the average wage of routine workers is higher than for manual non-routine workers. Schooling shares are included to illustrate the relationship between skills and occupations. Workers in routine occupations are on average better educated than manual non-routine workers, but have lower levels of education than cognitive non-routine workers. More details of the occupation classification are provided in Data Appendix E.

<table>
<thead>
<tr>
<th>Three-group classification</th>
<th>Contained occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive non-routine</td>
<td>Managers</td>
</tr>
<tr>
<td></td>
<td>Professionals</td>
</tr>
<tr>
<td></td>
<td>Technicians</td>
</tr>
<tr>
<td>Routine</td>
<td>Sales</td>
</tr>
<tr>
<td></td>
<td>Office and admin</td>
</tr>
<tr>
<td></td>
<td>Production, craft and repair</td>
</tr>
<tr>
<td></td>
<td>Operators, fabricators and laborers</td>
</tr>
<tr>
<td>Manual non-routine</td>
<td>Protective service</td>
</tr>
<tr>
<td></td>
<td>Food prep, buildings and grounds, cleaning</td>
</tr>
<tr>
<td></td>
<td>Personal care and personal services</td>
</tr>
</tbody>
</table>

Table 1: Occupation classification based on Acemoglu and Autor (2011)
Table 2: Summary Statistics for CPS Sample

<table>
<thead>
<tr>
<th></th>
<th>man. non-rout.</th>
<th>rout.</th>
<th>cogn. non-rout.</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment share in pct</td>
<td>10.98</td>
<td>58.68</td>
<td>30.34</td>
<td>100</td>
</tr>
<tr>
<td>Wage in 2016-$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11.89</td>
<td>16.91</td>
<td>25.36</td>
<td>18.98</td>
</tr>
<tr>
<td>St. dev.</td>
<td>6.15</td>
<td>9.31</td>
<td>14.24</td>
<td>11.74</td>
</tr>
<tr>
<td>Schooling Shares in pct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>24.73</td>
<td>13.63</td>
<td>1.3</td>
<td>11.11</td>
</tr>
<tr>
<td>High school</td>
<td>44.78</td>
<td>46.15</td>
<td>13.53</td>
<td>36.1</td>
</tr>
<tr>
<td>Some college</td>
<td>19.31</td>
<td>21.72</td>
<td>14.31</td>
<td>19.21</td>
</tr>
<tr>
<td>College</td>
<td>10.37</td>
<td>16.91</td>
<td>46.42</td>
<td>25.15</td>
</tr>
<tr>
<td>More than college</td>
<td>0.82</td>
<td>1.58</td>
<td>24.45</td>
<td>8.44</td>
</tr>
</tbody>
</table>

Note: Based on data from the NBER CPS Merged Outgoing Rotation Groups. The sample is explained in Section E.1. The classification into three categories is explained in Section E.2. Data is for the year 1993. Wages are in 2016-Dollars.

Table 3: Calibration - Skill-distribution parameters and skill-prices

<table>
<thead>
<tr>
<th>$\sigma_M$</th>
<th>$\sigma_R$</th>
<th>$\sigma_C$</th>
<th>$\rho_{MR}$</th>
<th>$\rho_{MC}$</th>
<th>$\rho_{RC}$</th>
<th>$Y_M$</th>
<th>$Y_R$</th>
<th>$Y_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.73</td>
<td>0.56</td>
<td>0.68</td>
<td>-0.68</td>
<td>-0.77</td>
<td>0.95</td>
<td>3.29</td>
<td>14.71</td>
<td>13.15</td>
</tr>
</tbody>
</table>

Note: Parameters obtained from ML estimation based on (160) using CPS MORG data for 1993. See Section E for a discussion of the data.

5.1.2 Estimation

I estimate the parameters of (48) by Maximum Likelihood. The results are given in Table 3. To assess the goodness of fit, I compute the estimated density of log wages for each occupation and plot it against a histogram of the data in Figure 1. Both line up well, and the model thus generates a realistic wage distribution. The plots also reveal considerable overlap of wage distributions across occupations.

5.2 Production function

I model the production function similar to Autor and Dorn (2013) who study the role of ICT capital for wage polarization. Like this paper, they distinguish between manual non-routine, routine and cognitive non-routine labor. Their production function has the feature that ICT capital substitutes for routine labor and complements non-routine labor, thereby affecting wages in the three occupations differentially. However, whereas in Autor and Dorn (2013) production of services and goods takes place in separate sectors, in my model a single consumption good

25 A log-Normal distribution is a good description of most of the wage distribution. However, wages at the top are better described by a Pareto distribution. In optimal tax analysis it is therefore common practice to append a Pareto tail to the top of the distribution. While this is crucial for the analysis of income taxes at the top, it is unlikely to change results for the optimal robot tax, which is the focus of this paper. Since appending a Pareto tail to the skill distribution would further complicate the estimation, this paper focuses on the log-Normal distribution.

26 Autor and Dorn (2013) refer to manual non-routine as “Services” and to cognitive non-routine as “Abstract”.

22
Figure 1: Goodness of fit: Estimated density vs. data

is produced by combining all factors according to

\[ Y(L, B) = A \, L_M^{\alpha} L_C^{\beta} \left( \mu L_R^{\rho - 1} + (1 - \mu) B^{\rho - 1} \right)^{\frac{\rho - 1}{\rho - 1 - (\alpha + \beta)}}. \]  

(50)

The inputs are manual non-routine labor \( L_M \), routine labor \( L_R \), cognitive non-routine labor \( L_C \), and robots \( B \). Parameter \( A \) is a Hicks-neutral productivity shifter, \( \alpha \) determines the income share going to manual labor, \( \beta \) governs the share of income that goes to cognitive labor, and \( 0 \leq \mu \leq 1 \) is the weight given to routine labor as compared to robots. Finally, \( \rho \) is the elasticity of substitution between robots and routine labor. Ceteris paribus, if \( \rho > 1 \), an increase in the amount of robots leads the wage rate of routine labor to fall relative to non-routine labor. Moreover, if \( \rho (\alpha + \beta) > 1 \), the wage rate for routine workers drops in absolute terms as the number of robots increases. The partial-equilibrium effect of robots on relative wage rates is summarized by the elasticities

\[ \varepsilon_{Y_M/Y_R,B} = \varepsilon_{Y_C/Y_R,B} = \frac{(1 - \mu) (\rho - 1) L_R^{1/\rho} B}{\rho \left( 1 - \mu \right) L_R^{1/\rho} B + \mu L_R B^{1/\rho}}, \]  

(51)

which enter the expression for the optimal robot tax. With \( \rho > 1 \), we have \( \varepsilon_{Y_M/Y_R,B} > 0 \) and \( \varepsilon_{Y_C/Y_R,B} > 0 \), and the production function (50) satisfies Assumption 1. The production function implies that in partial equilibrium, wage rates for manual non-routine labor and cognitive non-routine labor are symmetrically affected by a marginal change in robots. The same effect of robots on wage rates would result if one instead considered a production structure and preferences as in Autor and Dorn (2013), while setting the elasticity of substitution between goods and services in consumption to unity. In that case, the effect of robots on the price for manual non-routine labor would run via consumer demand, rather than via technology directly. Having specified the production technology, I proceed with calibrating aggregate labor inputs. To do so, I need to specify preferences and individual labor supply as well as the tax system.

---

27 In Autor and Dorn (2013) output of goods and services is combined on the consumption side. In contrast, one can interpret (50) as combining output of different sectors on the production side – without having to consider two commodities while generating the same pattern of factor-price responses. Cortes et al. (2017) use a similar production function as this paper when studying the role of automation for disappearing routine jobs.

28 Note that these conditions hold in a partial equilibrium setting. In the model, the choice of robots is endogenous, as is labor supply and occupational choice. Still, the conditions give an indication for the ranges of parameters required to generate certain wage responses.

29 Herrendorf et al. (2013) find that a consumption elasticity of substitution between manufacturing and services of close to one is in line with the structural transformation observed in the US.
Table 4: Calibration - Parameters for labor supply equation (57)

<table>
<thead>
<tr>
<th>Param.</th>
<th>Value</th>
<th>Explanation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>0.3</td>
<td>Labor Supply Elasticity</td>
<td>Blundell and Macurdy (1999); Meghir and Phillips (2010)</td>
</tr>
<tr>
<td>$t$</td>
<td>0.181</td>
<td>Tax Progressivity</td>
<td>Heathcote et al. (2017)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>6.56</td>
<td>Revenue Parameter</td>
<td>Inferred from Heathcote et al. (2017)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>1494</td>
<td>Income Scaling Parameter</td>
<td>Computed based on median income</td>
</tr>
</tbody>
</table>

5.3 Preferences and labor supply

Preferences over consumption and labor supply are quasi-linear and given by

$$U(c, \ell) = c - \frac{\ell^{1+\varepsilon}}{1+\varepsilon},$$

(52)

where $c$ is consumption, $\ell$ is individual labor supply and $\varepsilon$ is the labor supply elasticity. Based on evidence reported by Blundell and Macurdy (1999) and Meghir and Phillips (2010), I set $\varepsilon = 0.3$. Let $T'(y)$ denote the marginal income tax rate at income $y$. Using that preferences are quasi-linear, optimal labor supply $\ell$ is then implicitly given by the solution to

$$\ell(w)^{\frac{1}{\varepsilon}} = (1 - T'(y(w))) w.$$  

(53)

In the calibration, I use the parametric tax function proposed by Heathcote et al. (2017) as an approximation to the US income tax schedule, with

$$T(y) = y - \lambda y^{-t},$$  

(54)

where $t$ is referred to as the progressivity parameter, while $\lambda$ can be used to calibrate total tax revenue. Based on (54), marginal tax rates are given by

$$T'(y) = 1 - (1 - t) \lambda y^{-t}.$$  

(55)

Using that $y = \ell w$ and substituting (55) in (53), one can explicitly solve for

$$\ell(w) = \left[ (1 - t) \lambda w^{1-t} \right]^{\frac{1}{1-t\varepsilon}}.$$  

(56)

Heathcote et al. (2017) estimate $t = 0.181$ based on data from the PSID for the years between 2000 and 2006, and using the NBER TAXSIM (Feenberg and Coutts, 1993). The value of $\lambda$ is not reported in their paper, but can be inferred from their Figure 1 which features a marginal tax rate of 50% at an income of 500000 USD, implying $\lambda \approx 6.56$. Finally, the level of income generated by my model does not coincide with annual income. To have the appropriate marginal tax rates apply to incomes in my model, I introduce a scaling parameter $\gamma$ such that the median income in my model multiplied by $\gamma$ corresponds to the median income in the data, which implies $\gamma \approx 1494$. With the scaling factor $\gamma$, I now have

$$\ell(w) = \left[ (1 - t) \gamma^{-t} w^{1-t} \right]^{\frac{1}{1-t\varepsilon}}.$$  

(57)

All parameters for computing labor supply are summarized in Table 4.

Having calibrated aggregate labor supplies $L_i$, it remains to calibrate the parameters of the production function and to set the amount of robots. I target wage rates $Y_M$, $Y_R$ and $Y_C$ for the model to generate a realistic wage distribution.

---

30 The median income is roughly 500008.
5.4 Response to robots

I now turn to the external targets which primarily govern the model’s response to a change in the amount of robots, as well as the response to a tax on robots. The empirical evidence regarding the impact of robots on wages is still scarce. Using data from the International Federation of Robotics (IFR, 2014), Acemoglu and Restrepo (2017) estimate the effect of a change in exogenous exposure to industrial robots between 1993 and 2007 on wages and employment at the level of US commuting zones. Based on local estimates, they argue that at the level of the aggregate economy, an additional robot per thousand workers reduces average wages by 0.5%. Rather than interpreting this semi-elasticity as the impact of robots on average wages across all occupations, I consider it as the impact of robots on the average wage of workers in routine occupations.31 This interpretation is not inconsistent with the findings in Acemoglu and Restrepo (2017). At the commuting zone level, Acemoglu and Restrepo (2017) estimate that employment falls in occupations which I classify as routine, while employment in other occupations is hardly affected. This suggests that the negative impact of robots on wages may also be concentrated in routine occupations.32 To convert the semi-elasticity from Acemoglu and Restrepo (2017) into an elasticity, I use that in 1993 in the US there were 0.36 robots per thousand workers, implying an elasticity of robots on the average wage of routine workers of -0.0018.33 To target this elasticity in my model, I need to generate an exogenous increase in the number of robots. To do so, I exogenously lower the producer price of robots. As a result, the firm in the model adopts more robots, which impacts wages.

I also target the impact of a change in the price of robots on robot adoption. This is an important moment since a tax on robots affects robot adoption via the same channel: by changing the (consumer) price of robots. According to data from the IFR (2006), the quality adjusted price of robots dropped by about 60% between 1993 and 2005 (the latest date for which I have quality adjusted price data), while the stock of robots roughly tripled. For lack of better evidence, I treat the price change in the data as exogenous and as the only driver of robot adoption, implying a price elasticity of robot adoption of 5. Finally, for given parameters of the production function, the unit in which robots are measured determines the return to robots. However, there is no clear target for the return to robots. Although the IFR computes a price index for robots, it is only useful for assessing relative price changes. Instead of taking the initial amount of robots from the data, I therefore calibrate it such that a one-percent drop in the price of robots is in line with the targeted elasticities.

All targets, internal and external, are matched perfectly. I report the calibrated parameters in Table 5. The value of the productivity shifter $A$ has no clear interpretation. It is required to match levels of wage rates. The share in total income going to manual workers is 6%, whereas 45% goes to cognitive workers. The weight given to routine workers as compared to robots is close to 1. Finally, the elasticity of substitution between robots and routine work is 4.9. The high elasticity is required in order to generate the negative wage effect for routine workers.

Table 6 summarizes the elasticities for relative wage rates and average wages which are implied by the calibration, both for the model without occupational choice, as well as for a model

---

31In my model, an increase in the number of robots always increases the average wage. To see this, note that an increase in robots raises total output. Since next to robots, labor is the only production factor, and the production technology exhibits constant returns to scale, labor gains on aggregate.

32While Acemoglu and Restrepo (2017) report the impact of robots on wages by education category, they do not report results by occupation.

33To compute the optimal tax on robots, Costinot and Werning (2018b) also calculate an elasticity based on Acemoglu and Restrepo (2017). In contrast to this paper, their elasticity is an average over the elasticities of relative wages at the deciles of the wage distribution. Moreover, to convert semi-elasticities into elasticities, they use the number of robots in 2007 rather than 1993.
Table 5: Calibration - Production function parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>A</th>
<th>α</th>
<th>β</th>
<th>μ</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>30.68</td>
<td>0.06</td>
<td>0.45</td>
<td>0.99</td>
<td>4.90</td>
</tr>
</tbody>
</table>

Note: Parameters obtained from calibration to match the following targets: wage rates $Y_M$, $Y_R$, $Y_C$, the elasticity of the average routine wage wrt. robots, the elasticity of robot adoption wrt. the price of robots, and a drop in robot prices consistent with the elasticities.

Table 6: Calibration Results: Implied Elasticities

<table>
<thead>
<tr>
<th>Occ. Choice</th>
<th>$\varepsilon_{YM/YR,B}^{GE}$</th>
<th>$\varepsilon_{YC/YR,B}^{GE}$</th>
<th>$\varepsilon_{\bar{w}_M,B}^{GE}$</th>
<th>$\varepsilon_{\bar{w}_R,B}^{GE}$</th>
<th>$\varepsilon_{\bar{w}_C,B}^{GE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0.0052</td>
<td>0.0052</td>
<td>0.0007</td>
<td>0.0034</td>
<td>-0.0018</td>
</tr>
<tr>
<td>Yes</td>
<td>0.0029</td>
<td>0.0050</td>
<td>0.0007</td>
<td>0.0001</td>
<td>-0.0012</td>
</tr>
</tbody>
</table>

Note: $\varepsilon_{YM/YR,B}^{GE}$ and $\varepsilon_{YC/YR,B}^{GE}$ are the general-equilibrium (GE) equivalents of $\varepsilon_{YM/YR,B}$ and $\varepsilon_{YC/YR,B}$. $\varepsilon_{\bar{w}_i,B}^{GE}$ is the GE elasticity of the average wage wrt. robots. $\varepsilon_{\bar{w}_M,B}$ is the GE elasticity of the average wage in occupation $i \in \{M,R,C\}$ wrt. robots. Elasticities are computed by lowering the price of robots by 1% and by then computing the general-equilibrium response of robots and average wages given the calibrated parameters, and given a parametric tax system. Without occupational choice, general and partial equilibrium elasticities coincide.

with occupational choice. As in the calibration, elasticities with respect to the amount of robots are computed by exogenously lowering the price of robots by 1%, and by then using the implied equilibrium responses. As discussed above, in partial equilibrium, a change in robots affects wage rates for manual and cognitive labor symmetrically, and thus $\varepsilon_{YM/YR,B} = \varepsilon_{YC/YR,B}$. With fixed occupations, general-equilibrium elasticities and partial equilibrium elasticities coincide. Moreover, changes in average wages are driven by changes in wage rates only, and we therefore observe the same effect of robots on average wages for manual and cognitive workers, $\bar{w}_M$ and $\bar{w}_C$. The average wage of routine workers, $\bar{w}_R$, falls, as required to match the targeted elasticity. The effect of robots on the overall average is slightly positive. If occupational choice is possible, wages respond generally less to a change in robots. Routine workers who see their wages drop due to more robots being used, switch to other occupations – in this case mostly manual non-routine occupations – which dampens both the decline of routine wages and the increase of non-routine wages.

5.5 Social Welfare Weights

Before I can compute optimal policy, welfare weights need to be specified. I follow Rothschild and Scheuer (2013) and assume relative social welfare weights according to

$$\Psi (w) = 1 - (1 - F)^r,$$

where $r$ parametrizes the government’s desire to redistribute. Here, $r = 1$ corresponds to utilitarian preferences, which combined with quasi-linear utility would imply no redistribution. As $r \to \infty$, the welfare weights approach that of a Rawlsian social planner. I set $r = 1.3$ which leads to marginal tax rates in the range of those observed empirically.

---

34 The model with occupational choice is based on the same calibrated parameters as the model without occupational choice. The only difference is that individuals may switch occupation.

35 The average wage in occupation $i$ is given by $\bar{w}_i = \left( \int_{\bar{w}_L,B}^{\bar{w}_H,B} w f_{L,B}(w) \, dw \right)^{-1} \int_{\bar{w}_L,B}^{\bar{w}_H,B} w f_{L,B}(w) \, dw$.

36 This pattern is in line with Cortes et al. (2017), who find that those demographic groups which are associated with a decline in routine employment also account for an increase in employment in non-routine manual occupations (as well as for an increase in non-employment).
5.6 Optimal tax results

I compute optimal policy—non-linear income taxes and the robot tax—for two scenarios: short- and medium-run. In the short-run, individuals cannot switch occupation, whereas in the medium-run they can. At a given point in time, the two scenarios only differ by one factor: whether or not individuals respond to the introduction of a robot tax by switching occupation. To achieve this, I obtain results for the short-run as follows: first, equilibrium is computed for a model with occupational choice, but in which it is not possible to tax robots; next, I fix occupational choice, but allow for a tax on robots. In the short-run, the optimal tax on the stock of robots is 4.08%. It is thus optimal to distort the use of robots downward. What is the welfare gain of introducing a robot tax, provided that the income tax is already set optimally? I obtain the welfare impact by computing its consumption equivalent. The optimal robot tax is worth 0.00146% of GDP, which based on US per capita GDP in 2016 translates into 0.84$ per person per year. In contrast, in the medium-run with occupational choice, the optimal robot tax is 0.42%. Intuitively, individuals adjust their occupation in response to the wage changes brought about by the robot tax, thereby partly offsetting those changes. With occupational choice, the robot tax is thus less effective in compressing wages, and is therefore optimally smaller. Moreover, occupational choice lets the welfare gains evaporate: they are now reduced to a share of 0.00002% of GDP or 0.01$ per person per year.

Figure 2 plots optimal marginal income tax rates both for the case in which robots can and cannot be taxed. First, note that marginal tax rates follow the common U-shape as for example in Saez (2001). In the short-run, without occupational choice, the presence of a robot tax affects marginal taxes: tax rates are now higher at low to medium wages, and lower at high wages. For the simple model in Section 3, I derive that marginal income taxes in the presence of a robot tax are ceteris paribus lower for manual and cognitive workers. The effect on marginal income taxes for routine workers is ambiguous. As cognitive workers are concentrated at high wages, the result of lower taxes seems to carry over to the full model. Moreover, in terms of the simple
model, lower marginal tax rates for manual workers seem to be overturned by higher marginal taxes for routine workers. Of course, the equivalence with the simple model is imperfect as wage distributions now overlap occupations. If individuals can switch occupation, the ability to tax robots does not visibly change optimal marginal income taxes, which is not surprising, given that the optimal robot tax is very small.

Next, I ask how progress in robotics technology, expressed as a drop in the price of robots, affects the economy, as well as the optimal tax on robots. Again, I consider the short-run and the medium-run scenario. Figure 3 shows results for the short-run scenario in which employment shares are fixed at their initial level. The horizontal axis measures the drop in the robot price relative to its initial level, hence 0 corresponds to the initial price and 50 to a robot price which is 50% lower than its initial level. Without occupational choice, average wages are driven by changing wage rates alone, as the composition of skills within occupations remains constant. For a substantial range of robot prices, average wages hardly change. The reason is that the number of robots only increases slightly. It is only after the price of robots has dropped by about 60% that the number of robots increases strongly – and as a result average wages also change substantially. The average wage of routine workers decreases, whereas average wages for non-routine occupations increase in parallel. The optimal robot tax increases from around

---

Figure 3: Effect of robot price drop on employment and wages: without occupational choice

Note: The number of robots is per thousand workers.

---

One can also think of the horizontal axis as representing time. However, in that case the short-run and medium-run scenarios should simply be thought of as scenarios with fixed vs. flexible occupational choice, rather than representing different time horizons.
4% to 5% as robots get cheaper, but starts to fall once average wages of manual non-routine workers overtake those of routine workers. If individuals cannot switch occupations at all, the tax thus remains sizable for the most part. I now turn to the medium-run scenario.

The top row of Figure 4 shows the response of employment shares and average wages by occupation, assuming that occupational switching is possible. As robots get cheaper, they substitute for more and more routine workers, who then switch to either manual or cognitive occupations. Again, for a large range of robot prices, employment shares hardly change. Only after the robot price has dropped by about 60% do we observe substantial changes. As can be seen in the bottom row of Figure 4, this is the point at which the number of robots increases substantially, putting pressure on routine wages. As a result, individuals switch from routine occupations to non-routine occupations – and most notably to cognitive work. Recall, that wage rates for manual and cognitive workers are symmetrically affected by robots. The differential switching into non-routine occupations is thus driven by the skill distribution.

If one interprets the employment changes in the medium-run as changes in the cross-section, then they correspond well to the empirically documented employment polarization. For example, Acemoglu and Autor (2011) report that the employment share of cognitive non-routine workers increased by 4.6 percentage points between 1989 and 2007, whereas the share of manual non-routine workers increased by 3.5 percentage points. In contrast, the share of routine workers dropped by 8.1 percentage points. If, in contrast, one interprets the changes in employment shares in the medium-run as changes in the panel-dimension, the pattern is less in line with the data. In particular, Cortes et al. (2017) find no significant switching from routine occupations to cognitive non-routine occupations within individuals. However, they document switching from routine to manual non-routine occupations.

In contrast to the scenario without occupational choice, average wages now increase for all occupations, but less so for routine work than for non-routine work. There are two reasons for the increasing routine wage: First, as individuals move out of routine occupations, aggregate labor supply falls, which drives up the wage rate. Second, only those individuals who have a comparative advantage in routine occupations stay, which improves the skill composition. Still, eventually manual non-routine workers earn more on average than routine workers.

Comparing the number of robots across the two scenarios, we see that the use of robots increases much more strongly if occupational choice is possible. There are two reasons for this difference. First, with occupational choice, the allocation of skills is more efficient. As individuals move into occupations which are complementary with robots, the return to robot adoption is higher. Second, the robot tax is now substantially lower.

The robot tax remains close to constant for a large range of robot prices. Eventually it decreases to zero, increases again, and then even becomes slightly negative. Interestingly, this happens at a point at which routine workers who would benefit from a tax on robots are found at the bottom of the income distribution. In the simple model, a tax on robots would now unambiguously decrease inequality. However, with occupational choice and heterogeneity, effort reallocation and occupational shift effects need be considered. As a result, it is now optimal to stimulate employment in non-routine occupations for redistributive reasons. Still, the magnitude of the robot tax (or subsidy) in the presence of occupational choice remains very small – and approaches zero as the price of robots falls.

Guerreiro et al. (2017) also find that the robot tax goes to zero if the price of robots drops sufficiently. In their case, routine workers exit the labor market, and the remaining non-routine workers are homogeneously affected by robots. As a result, a tax on robots cannot compress wage differentials, and should therefore not be used. In this paper, instead of exiting the labor market, routine workers move into non-routine occupations. Since robots affect wages in both non-routine occupations in the same way, taxing robots can not anymore achieve wage

---

38 Based on Table 3a, where I consider professional, managerial and technical workers as cognitive non-routine, service workers as manual non-routine, and clerical, sales, and production workers as well as operators as routine.
compression, and robots should no longer be taxed.

5.7 Robustness

In this section, I discuss how strongly the results depend on the response of routine wages to an increase in robots. While the baseline calibration makes use of the best empirical evidence available, there is arguably still a lot of uncertainty regarding the effect of robots on wages. Since I consider this calibration target as both important and the least certain, I investigate how results change if the targeted elasticity is different. In particular, I ask how the results would change if the response of routine wages to robots were 10, 20 and 30 times as large as in the baseline. All other calibration targets are kept the same, for the model to generate the same income distribution as in the baseline version, and for taxes to have the same effect on firms’ choice of robots. I report the results in Table 7 in the Appendix. As before, I compute the optimal robot tax both for the short-run without occupational choice, and for the medium-run with occupational choice. I find that robot taxes are lower, the more responsive wages are to a change in robots.

In contrast, Costinot and Werning (2018b) find that the robot tax increases, ceteris paribus, in the elasticity of relative wages with respect to robots. However, the ceteris paribus condition is not satisfied in the robustness computations. In particular, keeping the income distribution and the price elasticity of robot adoption the same, while increasing the elasticity of routine

Figure 4: Effect of robot price drop on employment, wages and the robot tax: with occupational choice

Note: The number of robots is per thousand workers.
wages with respect to robots, requires robots to account for a larger share in total income ($\mu$ is lower at higher $\varepsilon_{\pi R,B}$).

Taking the changing income share of robots into account, my findings are not inconsistent with Costinot and Werning (2018b). According to their sufficient statistics formula, the optimal robot tax decreases in the share of income which goes to robots. Moreover, if the relative increase in the income share of robots is larger than the relative increase in the elasticity of relative wages with respect to robots, the optimal robot tax falls.

I also compute welfare for both, the short- and medium-run. In the short-run, the welfare benefits per capita now increase to around 20$ in the most elastic case. However, the medium-run welfare impact remains in the order of cents per person per year. The finding that in the medium-run, a tax on robots has a negligible effect on welfare is thus robust, even if robots have a much larger impact on wages than the available empirical evidence suggests. If at all, a tax on robots might thus only be worth considering in the short-run.

6 Conclusion

This paper studies the optimal taxation of robots and labor income in a model in which robots substitute for routine labor and complement non-routine labor. Intuition is developed in a stylized model based on Stiglitz (1982), which features intensive-margin labor supply and endogenous wages, but in which types are discrete and occupations are fixed. The full model then introduces continuous wage distributions and occupational choice, building upon (Rothschild and Scheuer, 2013, 2014).

I find that in general, the optimal robot tax is not zero, thereby violating production efficiency (Diamond and Mirrlees, 1971). The robot tax exploits general-equilibrium effects to compress the wage distribution. As a consequence, income taxation becomes less distortionary – which allows for more redistribution overall, and increases welfare. Since workers in routine occupations are concentrated at medium incomes, the sign of the optimal robot tax is theoretically ambiguous. Taxing robots reduces inequality at high incomes, thereby locally lowering income-tax distortions of labor-supply; but it increases inequality at low incomes, and thus locally worsens labor-supply distortions.

To assess the optimal robot tax quantitatively, I calibrate the model to the US economy. The calibration matches the distribution of wages and employment across manual non-routine, routine, and cognitive non-routine occupations. Moreover, it is informed by the labor-market impact of industrial robots studied by Acemoglu and Restrepo (2017). Technology is modeled similar to Autor and Dorn (2013).

I compute optimal policy for a short-run scenario, in which occupations are fixed, and a medium-run scenario in which individuals can switch occupation in response to the introduction of a robot tax. In the short-run, the optimal tax on the stock of robots is in the order of 4% and the consumption-equivalent welfare gain from introducing the tax is about 0.84$ per person per year. In the medium-run, as individuals switch occupation in response to a tax on robots, its effectiveness in compressing wages is reduced. The optimal robot tax and the welfare impact are then much smaller: 0.4% and 0.01$. If the price of robots falls, the share of routine workers approaches zero, as they move into non-routine occupations. Taxing robots can then not anymore achieve wage compression. As a consequence, the optimal robot tax goes to zero, together with its welfare impact. The negligible welfare impact in the medium-run is robust, even if wages respond much more strongly to an increase in robots than suggested by the data. The benefits of a robot tax in the short-run increase more substantially with the wage response, but remain very modest overall.

In light of the small welfare gains from taxing robots, this paper does not provide a strong case for a robot tax. Additional costs cast doubt on the optimality of a robot tax in practice. For example, with a tax on robots come considerable administrative costs as machinery needs to
be classified into robots and non-robots. Moreover, I have abstracted from implications which a tax on robots would have in an open economy. As any tax on capital, a tax on robots could impact a firm’s location choice with additional implications for inequality and welfare.
Appendix

A Derivations for simple model

A.1 Optimal robot tax

Assuming that only the adjacent downward-binding incentive constraints are relevant, maximized social welfare is given by the Lagrangian

\[ \mathcal{L} = f_M \psi_M V_M + f_R \psi_R V_R + f_C \psi_C V_C + \mu_{\xi R} \left( V_C - U \left( c_R (V_R, \ell_R), \ell_R \frac{Y_R (L, B)}{Y_C (L, B)} \right) \right) + \mu_{\eta R} \left( V_R - U \left( c_M (V_M, \ell_M), \ell_M \frac{Y_M (L, B)}{Y_R (L, B)} \right) \right) + \mu \xi (f_M \ell_M - L_M) + \mu \xi_r (f_R \ell_R - L_R) + \mu \xi_C (f_C \ell_C - L_C) + \mu \left( \sum_{i \in \mathcal{I}} f_i \ell_i Y_i (L, B) + Y_B (L, B) B - \sum_{i \in \mathcal{I}} f_i c_i - q B \right), \]  

with \( \mathcal{I} = \{ M, R, C \} \), where \( \mu \) is the multiplier on the resource constraint, \( \mu_{\xi R} \) is the multiplier on the incentive constraint for cognitive workers, and \( \mu_{\eta R} \) is the multiplier on the incentive constraint for routine workers. Moreover, \( \mu \xi_i \) is the multiplier on the consistency condition for occupation \( i \). To find an expression for the optimal robot tax, first write the resource constraint as \( Y (L, B) - \sum_{i \in \mathcal{I}} f_i c_i - q B = 0 \). Then differentiate the Lagrangian with respect to \( B \) to obtain

\[ \frac{\partial \mathcal{L}}{\partial B} = -\mu_{\xi R} U \left( c_R (V_R, \ell_R), \ell_R \frac{Y_R (L, B)}{Y_C (L, B)} \right) \frac{\partial }{\partial B} \left( Y_R (L, B) \right) - \mu_{\eta R} U \left( c_M (V_M, \ell_M), \ell_M \frac{Y_M (L, B)}{Y_R (L, B)} \right) \frac{\partial }{\partial B} \left( Y_M (L, B) \right) + \mu (\dot{Y}_B (L, B) - q). \]  

Define elasticities of relative wage rates with respect to robots as

\[ \varepsilon_{Y_C / Y_R, B} = -\varepsilon_{Y_R / Y_C, B} \equiv \frac{\partial}{\partial B} \left( \frac{Y_R (L, B)}{Y_C (L, B)} \right) \frac{Y_C (L, B)}{Y_R (L, B)} B, \]  

\[ \varepsilon_{Y_M / Y_R, B} \equiv \frac{\partial}{\partial B} \left( \frac{Y_M (L, B)}{Y_R (L, B)} \right) \frac{Y_R (L, B)}{Y_M (L, B)} B. \]  

Use that

\[ Y_B (L, B) = p = (1 + \tau) q \leftrightarrow \dot{Y}_B (L, B) - q = \tau q, \]  

set (60) equal to zero, rearrange and divide by \( \mu \) to get

\[ \tau q B = -\eta_{CR} U \left( c_R (V_R, \ell_R), \ell_R \frac{Y_R (L, B)}{Y_C (L, B)} \right) \frac{\partial }{\partial B} \left( Y_R (L, B) \right) \varepsilon_{Y_C / Y_R, B} + \eta_{RM} U \left( c_M (V_M, \ell_M), \ell_M \frac{Y_M (L, B)}{Y_R (L, B)} \right) \frac{\partial }{\partial B} \left( Y_M (L, B) \right) \varepsilon_{Y_M / Y_R, B}. \]  

Now define the incentive effects in a similar way as Rothschild and Scheuer (2013) (using that \( w_i = Y_i \) and suppressing some arguments)

\[ I_{CR} \equiv -\eta_{CR} U \left( c_R, \ell_R \frac{w_R}{w_C} \right) \ell_R \frac{w_R}{w_C}. \]
and

\[ I_{RM} \equiv -\eta_{RM} U_\ell \left( c_M, \ell_M \frac{w_M}{w_R} \right) \frac{\ell_M w_M}{w_R}, \tag{66} \]

to write

\[ \tau q B = \varepsilon_{Yc/Yr,B} I_{CR} - \varepsilon_{YM/Yr,B} I_{RM}. \tag{67} \]

A.1.1 Expressions for incentive effects

To obtain expressions for the incentive effects (65) and (66), one needs to derive expressions
for the multipliers \( \eta_{CR} \) and \( \eta_{RM} \). To do so, differentiate the Lagrangian in (59) with respect to
indirect utilities to obtain

\[ \frac{\partial L}{\partial V_M} = f_M \psi_M - \mu \eta_{RM} U_c \left( c_M, \ell_M \frac{w_M}{w_R} \right) \frac{\partial c_M}{\partial V_M} - \mu f_M \frac{\partial c_M}{\partial V_M}. \tag{68} \]

Equating to zero, using that \( \frac{\partial c_M}{\partial V_M} = \frac{1}{U_c(c_M, \ell_M)} \) and rearranging yields

\[ \eta_{RM} = f_M \left( \frac{1}{\psi_M} - \frac{1}{U_c(c_M, \ell_M)} \right) \frac{U_c(c_M, \ell_M)}{U_c\left( c_M, \ell_M \frac{w_M}{w_R} \right)}. \tag{69} \]

Analogously, we obtain

\[ \frac{\partial L}{\partial V_R} = f_R \psi_R - \mu \eta_{CR} U_c \left( c_R, \ell_R \frac{w_R}{w_C} \right) \frac{\partial c_R}{\partial V_R} + \mu \eta_{RM} - \mu f_R \frac{\partial c_R}{\partial V_R}, \tag{70} \]

which after equating to zero, substituting for \( \frac{\partial c_R}{\partial V_R} = \frac{1}{U_c(c_R, \ell_R)} \) and rearranging yields

\[ \eta_{RM} = f_R \left( \frac{1}{U_c(c_R, \ell_R)} - \frac{1}{\psi_R} \right) + \eta_{CR} U_c \left( c_R, \ell_R \frac{w_R}{w_C} \right) \frac{1}{U_c(c_R, \ell_R)}. \tag{71} \]

Finally, we have

\[ \frac{\partial L}{\partial V_C} = f_C \psi_C + \mu \eta_{CR} - \mu f_C \frac{\partial c_C}{\partial V_C}, \tag{72} \]

which after equating to zero, using \( \frac{\partial c_C}{\partial V_C} = \frac{1}{U_c(c_C, \ell_C)} \) and rearranging becomes

\[ \eta_{CR} = f_C \left( \frac{1}{U_c(c_C, \ell_C)} - \frac{1}{\psi_C} \right). \tag{73} \]

With quasi-linear utility as in (1) we have \( \mu = 1 \) and \( U_c = 1 \). As a result, one obtains

\[ \eta_{CR} = f_C (1 - \psi_C), \tag{74} \]

and

\[ \eta_{RM} = f_M (\psi_M - 1) = f_R (1 - \psi_R) + \eta_{CR} = f_R (1 - \psi_R) + f_C (1 - \psi_C). \tag{75} \]

Moreover, quasi-linear utility leads to \( U_\ell = -\ell^2 \) such that the incentive effects are

\[ I_{CR} = f_C (1 - \psi_C) \left( \ell_R \frac{w_R}{w_C} \right)^{1 + \frac{1}{2}}, \tag{76} \]

and

\[ I_{RM} = f_M (\psi_M - 1) \left( \ell_M \frac{w_M}{w_R} \right)^{1 + \frac{1}{2}}. \tag{77} \]

A.2 Optimal income tax

To derive expressions for the optimal marginal income tax rates, I differentiate the Lagrangian
(59) with respect to individual labor supplies \( \ell_i \).
A.2.1 Expression for $T'_M$

First consider $\partial L / \partial \ell_M$, using that $Y_i = w_i$ for $i \in I$ and suppressing some arguments

$$\frac{\partial L}{\partial \ell_M} = -\mu \eta_{RM} \left[ U_c \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \frac{\partial c(V_M, \ell_M)}{\partial \ell_M} + U_\ell \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \frac{w_M}{w_R} \right] + \mu \xi_M f_M + \mu f_M \left( w_M - \frac{\partial c}{\partial \ell_M} \right). \tag{78}$$

Use that

$$\frac{\partial c}{\partial \ell} = -\frac{U_\ell (c(V_i, \ell_i), \ell_i)}{U_c (c(V_i, \ell_i), \ell_i)} = w_i (1 - T'_i), \tag{79}$$

where the last step is based on the definition of marginal tax rates. Substituting in (78), I obtain

$$\frac{\partial L}{\partial \ell_M} = -\mu \eta_{RM} \left[ U_c \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) w_M (1 - T'_M) + U_\ell \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \frac{w_M}{w_R} \right] + \mu \xi_M f_M + \mu f_M \left( w_M - w_M (1 - T'_M) \right). \tag{80}$$

Setting equal to zero, dividing by $\mu$ and collecting terms yields

$$\frac{T'_M + \xi_M}{1 - T'_M} = \frac{1}{f_M} \eta_{RM} \left[ U_c \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) - \frac{U_c (c(V_M, \ell_M), \ell_M)}{U_\ell (c(V_M, \ell_M), \ell_M)} U_\ell \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \frac{w_M}{w_R} \right]. \tag{81}$$

With quasi-linear utility the expression becomes

$$\frac{T'_M + \xi_M}{1 - T'_M} = (\psi_M - 1) \left( 1 - \left( \frac{w_M}{w_R} \right)^{1+\frac{1}{\psi}} \right). \tag{82}$$

A.2.2 Expression for $T'_R$

The first-order condition with respect to $\ell_R$ is

$$\frac{\partial L}{\partial \ell_R} = -\mu \eta_{CR} \left[ U_c \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) \frac{\partial c}{\partial \ell_R} + U_\ell \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) \frac{w_R}{w_C} \right] + \mu \xi_R f_R + \mu f_R \left( w_R - \frac{\partial c}{\partial \ell_R} \right). \tag{83}$$

I thus obtain

$$\frac{T'_R + \xi_R}{1 - T'_R} = \frac{1}{f_R} \eta_{CR} \left[ U_c \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) - \frac{U_c (c(V_R, \ell_R), \ell_R)}{U_\ell (c(V_R, \ell_R), \ell_R)} U_\ell \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) \frac{w_R}{w_C} \right], \tag{84}$$

which with quasi-linear utility translates into

$$\frac{T'_R + \xi_R}{1 - T'_R} = \frac{f_C}{f_R} (1 - \psi_C) \left( 1 - \left( \frac{w_R}{w_C} \right)^{1+\frac{1}{\psi}} \right). \tag{85}$$
A.2.3 Expression for $T_C^\prime$

Finally, the first-order condition with respect to $\ell_C$ is

\[ \frac{\partial L}{\partial \ell_C} = \mu \xi_C f_C + \mu f_C \left( w_C - \frac{\partial c}{\partial \ell_C} \right). \]  

(86)

Setting equal to zero, using (79) and rearranging leads to

\[ T_C^\prime = -\frac{\xi_C}{Y_C}. \]  

(87)

A.2.4 Expressions for multipliers $\xi_i$

To derive expressions for the multipliers $\xi_i$, differentiate the Lagrangian (59) with respect to aggregate effective labor supplies $L_i$ to obtain

\[ \frac{\partial L}{\partial L_i} = -\mu \eta_{CR} U_\ell \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) \ell_R \frac{\partial}{\partial L_i} \left( \frac{w_R}{w_C} \right) \]

\[ -\mu \eta_{RM} U_\ell \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \ell_M \frac{\partial}{\partial L_i} \left( \frac{w_M}{w_R} \right) \]

\[ -\mu \xi_i + \mu \left( \sum_{j \in I} f_j \ell_j \frac{\partial Y_j}{\partial L_i} + \frac{\partial Y_B}{\partial L_i} B \right). \]  

(88)

Substitute $f_j \ell_j = L_j$. By Euler’s Theorem, the effect on the resource constraint is zero. Using that at the optimum, a change in $L_i$ has no effect on welfare and rearranging, I obtain

\[ \xi_i = -\eta_{CR} U_\ell \left( c(V_R, \ell_R), \ell_R \frac{w_R}{w_C} \right) \ell_R \frac{\partial}{\partial L_i} \left( \frac{w_R}{w_C} \right) \]

\[ -\eta_{RM} U_\ell \left( c(V_M, \ell_M), \ell_M \frac{w_M}{w_R} \right) \ell_M \frac{\partial}{\partial L_i} \left( \frac{w_M}{w_R} \right). \]  

(89)

Using the definitions of the incentive effects (76) and (77), we arrive at

\[ \xi_i = I_{CR} \frac{\partial}{\partial L_i} \left( \frac{w_R}{w_C} \right) \frac{w_C}{w_R} \]

\[ + I_{RM} \frac{\partial}{\partial L_i} \left( \frac{w_M}{w_R} \right) \frac{w_R}{w_M}. \]  

(90)

Now define the semi-elasticities of relative wages with respect to $L_i$ as

\[ \tilde{\varepsilon}_{w_R/w_C,L_i} = \frac{\partial}{\partial L_i} \left( \frac{w_R}{w_C} \right) \frac{w_C}{w_R} \]  

(91)

and

\[ \tilde{\varepsilon}_{w_M/w_R,L_i} = \frac{\partial}{\partial L_i} \left( \frac{w_M}{w_R} \right) \frac{w_R}{w_M} \]  

(92)

to write

\[ \xi_i = \tilde{\varepsilon}_{w_R/w_C,L_i} I_{CR} + \tilde{\varepsilon}_{w_M/w_R,L_i} I_{RM}. \]  

(93)
Signing the multipliers. The sign of the multiplier is determined by the terms $\frac{\partial}{\partial L_i} \left( \frac{w_M}{w_R} \right)$ and $\frac{\partial}{\partial L_i} \left( \frac{w_M}{w_C} \right)$. The sign of $\xi_i$ is unambiguous if both of these terms have the same sign. Consider $\xi_M$. We have $\frac{\partial}{\partial L_M} \left( \frac{w_M}{w_R} \right) < 0$. A sufficient condition for $\xi_M < 0$ is thus $\frac{\partial}{\partial L_M} \left( \frac{w_M}{w_C} \right) < 0$. Now consider $\xi_C$. We have $\frac{\partial}{\partial L_C} \left( \frac{w_R}{w_C} \right) > 0$, and hence $\frac{\partial}{\partial L_C} \left( \frac{w_M}{w_R} \right) > 0$ is a sufficient condition for $\xi_C > 0$. Finally, since $\frac{\partial}{\partial L_R} \left( \frac{w_R}{w_C} \right) < 0$, and $\frac{\partial}{\partial L_R} \left( \frac{w_M}{w_R} \right) > 0$ the sign of $\xi_R$ is ambiguous and depends on the magnitudes of the different terms in (93).

B Optimal tax on robots with continuous types and occupational choice

In order to derive an expression for the optimal tax on robots, I use that at the optimum, a marginal change in robots $B$, has no first-order welfare effect. The welfare effect of a marginal change in $B$ corresponds to differentiating the Lagrangian of the inner problem, evaluated at the optimal allocation, with respect to $B$. The Lagrangian of the inner problem is given by

$$
\mathcal{L} = \int_{\mathcal{L},B}^{\mathcal{L}} V(c(w), \ell(w)) d\Psi(w)
$$

$$
+ \mu \int_{\mathcal{L},B}^{\mathcal{L}} \eta(w) U_{\ell} c(V(w), \ell(w), \ell(w)) \frac{\ell(w)}{w} d\Psi(w)
$$

$$
+ \mu \xi_M \left( L_M - \frac{1}{Y_M (L, B)} \right) \int_{\mathcal{L},B}^{\mathcal{L}} w \ell(w) f_{L,B}^M (w) dw
$$

$$
+ \mu \xi_R \left( L_R - \frac{1}{Y_R (L, B)} \right) \int_{\mathcal{L},B}^{\mathcal{L}} w \ell(w) f_{L,B}^R (w) dw
$$

$$
+ \mu \xi_C \left( L_C - \frac{1}{Y_C (L, B)} \right) \int_{\mathcal{L},B}^{\mathcal{L}} w \ell(w) f_{L,B}^C (w) dw
$$

$$
+ \mu \left( \int_{\mathcal{L},B}^{\mathcal{L}} w \ell(w) f(w) dw + Y_B (L, B) B - \int_{\mathcal{L},B}^{\mathcal{L}} c(V(w), \ell(w)) f(w) dw - qB \right).
$$

The impact of a marginal change in $B$ can be decomposed into four effects. First, there is a direct effect on the resource constraint. The other three effects result from the impact which a change in $B$ has on wages: The second effect is the direct result from a change in wages, leading to changes in $\ell(w)$ and $V(w)$. The third effect is the indirect result of changing wages. As wages change, individuals move along the schedules $\ell(w)$ and $V(w)$. Finally, as relative wage rates change, some individuals switch occupation, which has an effect on the consistency conditions.

Instead of computing the effects by holding the schedules $\ell(w)$ and $V(w)$ fixed, using the envelope theorem, I follow Rothschild and Scheuer (2014) and construct variations in the schedules $\ell(w)$ and $V(w)$ which simplify the derivations. The idea is as follows: Instead of having to take into account that changes in $B$ alter the wage densities, the adjustment of $\ell(w)$ and $V(w)$ is chosen such that it offsets, at each $w$, changes which otherwise would require adjusting the densities. Denote the adjusted schedules by $\ell(w)$ and $V(w)$.

Schedule variations. In what follows, I first derive $\tilde{\ell}(w)$ and $\tilde{V}(w)$ which requires some preparation: Indicate occupations by $i \in \mathcal{I} \equiv \{M, R, C\}$ and denote by

$$
\beta^i_B (L, B) = \frac{\partial Y_i (L, B)}{\partial B} \frac{1}{Y_i (L, B)},
$$

(95)
the semi-elasticity of the skill-price $Y_i(L, B)$ with respect to $B$. I define the indicator $q^i_{L, B}(\theta)$ such that

$$ q^i_{L, B}(\theta) = \begin{cases} 1, & \text{if } \theta \text{ works in } i \\ 0, & \text{otherwise.} \end{cases} \tag{96} $$

Using that wages are given by

$$ w_{L, B}(\theta) = \max \{ Y_M(L, B) \theta_M, Y_R(L, B) \theta_R, Y_C(L, B) \theta_C \}, \tag{97} $$

the wage of individual $\theta$ can thus be written as

$$ w_{L, B}(\theta) = \sum_{i \in I} q^i_{L, B}(\theta) \theta Y_i(L, B). \tag{98} $$

The semi-elasticity of wages with respect to $B$ for individual $\theta$ is thus

$$ \frac{\partial w_{L, B}(\theta)}{\partial B} \frac{1}{w_{L, B}(\theta)} = \sum_{i \in I} q^i_{L, B}(\theta) \frac{\partial Y_i(L, B)}{\partial B} \frac{1}{Y_i(L, B)} = \sum_{i \in I} q^i_{L, B}(\theta) \beta^i_{B}(L, B). \tag{99} $$

Like Rothschild and Scheuer (2014), I now construct $\ell'(w)$ and $\tilde{V}(w)$ such that at each $w$, changes in average labor supply and indirect utility which come from individuals shifting along $\ell(w)$ and $V(w)$ are offset. I focus on deriving $\ell(w)$. The derivations for $\tilde{V}(w)$ are analogue. First, consider an individual $\theta$. A change $dB$, which leads to a change in $w_{L, B}(\theta)$, causes $\theta$ to adjust labor supply by

$$ \ell'(w) \frac{\partial w_{L, B}(\theta)}{\partial B} dB = \ell'(w) w \sum_{i \in I} q^i_{L, B}(\theta) \beta^i_{B}(L, B) dB, \tag{100} $$

where I use (99). Now, note that the same wage $w$ can be earned by different individuals if wage distributions overlap across occupations. In order to compute the average adjustment in labor supply, we need to compute the expected adjustment over all types $\theta$ earning $w$, that is

$$ \ell'(w) w \sum_{i \in I} E \left[ q^i_{L, B}(\theta) | w \right] \beta^i_{B}(L, B) dB = \ell'(w) w \sum_{i \in I} \frac{f^i_{L, B}(w)}{f_{L, B}(w)} \beta^i_{B}(L, B) dB, \tag{101} $$

where I use that $E \left[ q^i_{L, B}(\theta) | w \right]$ corresponds to the share of individuals who earn $w$ in occupation $i$, that is $f^i(w) / f(w)$. I now obtain the adjusted schedule $\tilde{\ell}(w)$ by subtracting the change in $\ell(w)$, induced by a change $dB$, from $\ell(w)$, that is

$$ \tilde{\ell}(w) \equiv \ell(w) - \ell'(w) w \delta^B_{L, B}(w) dB, $$

where I define

$$ \delta^B_{L, B}(w) \equiv \sum_{i \in I} \frac{f^i_{L, B}(w)}{f_{L, B}(w)} \beta^i_{B}(L, B). \tag{102} $$

The adjusted schedule $\tilde{V}(w)$ is

$$ \tilde{V}(w) = V(w) - V'(w) w \delta^B_{L, B}(w) dB. \tag{103} $$

Since $l(w)$ and $V(w)$ are chosen optimally, by the envelope theorem, a marginal adjustment to $\tilde{\ell}(w)$ and $\tilde{V}(w)$ brought about by a marginal change $dB$ has no first-order effect on welfare. I now consider the effect of a change $dB$ on the different parts of the Lagrangian. The objective is not affected by a change $dB$.\footnote{An exception is the case in which the planner redistributes based on non-welfarist principles. For example, the planner might favor redistribution to certain occupations based on criteria other than the distribution of indirect utilities. In this case, effects on the objective need to be taken into account, to which Rothschild and Scheuer (2013, 2014) refer to as redistributive effects.}

38
Incentive constraint effect. The incentive constraint is as in Rothschild and Schoneuer (2014), who show that the schedule modification to $\ell (w)$ and $\hat{V} (w)$ changes

$$V' (w) - U_\ell (c (w), \ell (w)) \ell (w) / w$$

by $-V' (w) w \, d \delta_{L,B} (w) / dw$. Integrating over all wages then leads to the following effect on the incentive constraint

$$- \sum_{i \in I} \beta_B^i (L, B) \mu \int_{\mathbb{R}^L,B} \eta (w) \frac{\ell (w)}{U_c (w)} V' (w) w \frac{d}{dw} \left( \frac{f_{L,B}^i (w)}{f_{L,B}(w)} \right) dw \, dB$$

$$= - \mu \sum_{i \in I} \beta_B^i (L, B) I_i (L, B) dB,$$

with the incentive effect

$$I_i (L, B) = \int_{\mathbb{R}^L,B} \eta (w) \frac{\ell (w)}{U_c (w)} V' (w) w \frac{d}{dw} \left( \frac{f_{L,B}^i (w)}{f_{L,B}(w)} \right) dw.$$

Resource constraint effect. The expression in (94) pertaining to the resource constraint is

$$\mu \left( \int_{\mathbb{R}^L,B} w \ell (w) f (w) dw - \int_{\mathbb{R}^L,B} c (V (w), \ell (w)) f (w) dw + Y_B (L, B) B - q B \right).$$

First, a change in $B$ has a direct effect on $Y_B (L, B) B - q B$, given by

$$\frac{\partial Y_B (L, B)}{\partial B} B + Y_B (L, B) - q.$$

Second, there is a direct effect on $w$ in the first integrand, leading to

$$\int_{\mathbb{R}^L,B} \delta_{L,B}^i (w) w \ell (w) f (w) dw.$$

Next, there are direct effects on $\ell (w)$ and $V (w)$, and thus on $c (V (w), \ell (w))$. However, these effects are exactly canceled out by varying the schedules to $\ell (w)$ and $V (w)$ to offset the indirect effect. Use $\delta_{L,B}^i (w) = \sum_{i \in I} \beta_B^i (L, B) f_{L,B} (w)$ to rewrite (109) as

$$\int_{\mathbb{R}^L,B} \delta_{B}^i (w) w \ell (w) f (w) dw$$

$$= \sum_{i \in I} \beta_B^i (L, B) \int_{\mathbb{R}^L,B} w \ell (w) f_{L,B}^i (w) dw$$

$$= \sum_{i \in I} \frac{\partial Y_i (L, B)}{\partial B} \frac{1}{Y_i (L, B)} \int_{\mathbb{R}^L,B} w \ell (w) f_{L,B}^i (w) dw$$

$$= \sum_{i \in I} \frac{\partial Y_i (L, B)}{\partial B} L_i.$$

Now use that due to linear homogeneity of the production function,

$$\sum_{i \in I} \frac{\partial Y_i (L, B)}{\partial B} L_i + \frac{\partial Y_B (L, B)}{\partial B} B = 0,$$

hence adding (108) and (109) and multiplying by $\mu$ yields the resource constraint effect

$$\mu (Y_B (L, B) - q).$$

(112)
Consistency condition effects. Next, I turn to the effects on the consistency conditions. There are effort-reallocation effects and occupational shift effects. Consider the consistency condition for $M$. The derivations for the other consistency conditions are analogue.

Effort-reallocation effect. First, rather than writing the condition in terms of wages, write it in terms of types $\theta$ as

$$L_M - \int_\Theta \theta_M \ell_M (\theta) dF (\theta).$$

(113)

Now use that in the Roy model individuals fully specialize and write

$$\ell_M (\theta) = \ell (w) q^M_{L,B} (\theta),$$

(114)

where $w = Y_M (L, B) \theta_M$. The integrand can then be written as

$$\theta_M q^M_{L,B} (\theta) \ell (w).$$

(115)

A change in $B$ affects the expression via three channels: First, there is a direct effect on $\ell (w)$ as a change in $B$ affects wages. Second, there is an indirect effect, as due to a change in wages, individuals move along the schedule $\ell (w)$. Third, a change in $B$ affects relative wage rates across sectors, and thus occupational choice, captured by $q^M_{L,B} (\theta)$. Here, I focus on the first two effects. The third effect will be discussed as occupational shift effect below.

For a single individual $\theta$, the direct effect changes (115) by

$$\theta_M q^M_{L,B} (\theta) \ell' (w) w \sum_{i \in I} \hat{q}^i_{L,B} (\theta) \beta^i_B (L, B) dB$$

(116)

$$= \frac{1}{Y_M (L, B)} \sum_{i \in I} \beta^i_B (L, B) \ell' (w) w^2 q^M_{L,B} (\theta) \hat{q}^i_{L,B} (\theta) dB,$$

where the second step substituted $\theta_M = \frac{w}{M (L, B)}$ and rearranged. To compute the effect at $w = Y_M (L, B) \theta_M$, one needs to take the expectation over all individuals $\theta$ who earn $w$, leading to

$$\frac{1}{Y_M (L, B)} \sum_{i \in I} \beta^i_B (L, B) \ell' (w) w^2 E [q^M_{L,B} (\theta) \hat{q}^i_{L,B} (\theta) | w] dB.$$  

(117)

To offset the indirect effect on (115), change the schedule $\ell (w)$ to $\tilde{\ell} (w)$ by subtracting

$$\ell' (w) w \sum_{i \in I} E [q^i_{L,B} (\theta) | w] \beta^i_B (L, B) dB$$

(118)

from $\ell (w)$, which changes (115) by

$$- \theta_M q^M_{L,B} (\theta) \ell' (w) w \sum_{j \in I} E [q^j_{L,B} (\theta) | w] \beta^j_B (L, B) dB.$$  

(119)

Again, this expression is for a single individual. To compute the effect at $w$, take the expectation over all $\theta$ earning $w$, which by the law of iterated expectations yields

$$- \theta_M \ell' (w) w \sum_{j \in I} E [q^j_{L,B} (\theta) | w] \beta^j_B (L, B) dB$$

(120)

$$= - \frac{1}{Y_M (L, B)} \sum_{i \in I} \beta^i_B (L, B) \ell' (w) w^2 E [q^M_{L,B} (\theta) | w] E [\hat{q}^i_{L,B} (\theta) | w] dB.$$
Combine (117) and (120) to arrive at the change in (115) due to the direct and indirect effect

$$\frac{1}{Y_M(L, B)} \sum_{i \in I} \beta_i^j (L, B) \ell' (w) w^2 \times$$

$$\left( E \left[ q_{L,B}^M (\theta) q_{L,B}^j (\theta) | w \right] - E \left[ q_{L,B}^M (\theta) | w \right] E \left[ q_{L,B}^j (\theta) | w \right] \right) dB$$

(121)

Integrate over wages to obtain

$$\sum_{i \in I} \beta_i^j (L, B) \frac{1}{Y_M(L, B)} \int_{w_{L,B}}^{\bar{w}_{L,B}} \ell' (w) w^2 \text{Cov} \left[ q_{L,B}^i (\theta), q_{L,B}^j (\theta) | w \right] f (w) dw \ dB,$$

and define

$$C_{iM} = \frac{1}{Y_M(L, B)} \int_{w_{L,B}}^{\bar{w}_{L,B}} \ell' (w) w^2 \text{Cov} \left[ q_{L,B}^i (\theta), q_{L,B}^M (\theta) | w \right] f (w) dw,$$

(123)

and generally for \( j \in I \)

$$C_{ij} = \frac{1}{Y_j(L, B)} \int_{w_{L,B}}^{\bar{w}_{L,B}} \ell' (w) w^2 \text{Cov} \left[ q_{L,B}^i (\theta), q_{L,B}^j (\theta) | w \right] f (w) dw.$$ (124)

The effort-reallocation effect for the consistency condition which corresponds to occupation \( j \in I \) is then

$$- \mu \xi_j \sum_{i \in I} \beta_i^j (L, B) C_{ij} \ dB.$$ (125)

**No effort-reallocation effect if wage distributions do not overlap.** Now suppose that wage distributions do not overlap sectors, that is, \( q_{L,B}^i (\theta) | w = 0 \) for \( i \neq M \). The expression in (121) then becomes

$$\frac{1}{Y_M(L, B)} q_{B}^M (L, B) \ell' (w) w^2 \times$$

$$\left( E \left[ q_{L,B}^M (\theta) q_{L,B}^M (\theta) | w \right] - E \left[ q_{L,B}^M (\theta) | w \right] E \left[ q_{L,B}^M (\theta) | w \right] \right) dB.$$ (126)

Now use that \( q_{L,B}^M (\theta) q_{L,B}^M (\theta) = q_{L,B}^M (\theta) \). Moreover, with no overlap of distributions all individuals who earn \( w = Y_M(L, B) \) \( \theta \) are in occupation \( M \), and thus \( q_{L,B}^M (\theta) | w = 1 \). As a result, the term in parenthesis becomes

$$E \left[ q_{L,B}^M (\theta) q_{L,B}^M (\theta) | w \right] - E \left[ q_{L,B}^M (\theta) | w \right] E \left[ q_{L,B}^M (\theta) | w \right]$$

$$= E \left[ q_{L,B}^M (\theta) | w \right] - E \left[ q_{L,B}^M (\theta) | w \right] E \left[ q_{L,B}^M (\theta) | w \right]$$ (127)

$$= 1 - 1,$$

and hence there is no effort-reallocation effect if wage distributions do not overlap across occupations.

**Occupational-shift effect.** To derive the impact of occupational change on the consistency conditions, I focus on the effect on the condition for occupation \( M \). The derivations for the other consistency conditions are analogue. Instead of writing the consistency conditions in terms of types \( \theta \), now write them again in terms of wages. The condition for occupation \( M \) is thus

$$L_M - \int_{w_{L,B}}^{\bar{w}_{L,B}} \frac{1}{Y_M(L, B)} w \ell' (w) f (w) dw = 0.$$ (128)
Focus on the effect on \( \frac{1}{Y_M(L,B)} y(w) \), with income \( y(w) \equiv w \ell(w) \). I first derive how income \( y(w) \) earned in occupation \( M \) changes due to occupational shifts in response to an increase in \( B \). A change in \( B \) alters wage rates \( Y_i(L,B) \), which in turn affect occupational choice.

Write the impact of a marginal increase in \( B \) on wage rate \( Y_i(L,B) \) as \( Y_i(L,B) \beta_B(L,B) \).

Now, first consider how a marginal increase in \( Y_R(L,B) \) affects occupational choice, and thus income \( y(w) \) earned in occupation \( M \). As \( Y_R(L,B) \) increases, individuals are going to shift from occupation \( M \) to \( R \). Since I consider a marginal change in \( Y_R \), I focus on those individuals who are indifferent between \( M \) and \( R \), which implies that they earn the same wage in both occupations, and thus

\[
\theta_M Y_M(L,B) = \theta_R Y_R(L,B) \Leftrightarrow \theta_R = \frac{\theta_M Y_M(L,B)}{Y_R(L,B)}.
\] (129)

Moreover, individuals need to be better off working in occupation \( M \) than working in occupation \( C \), thus

\[
\theta_M Y_M(L,B) \geq \theta_C Y_C(L,B) \Leftrightarrow \theta_C \leq \frac{\theta_M Y_M(L,B)}{Y_C(L,B)}.
\] (130)

Having characterized the affected individuals, I next, consider how a change in relative prices due to a change in \( Y_R \), \( \Delta Y_R \), affects conditions (129) and (130). We get

\[
\theta_R = \theta_M \frac{Y_M(L,B)}{Y_R(L,B)} + \theta_M \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L,B)}{Y_R(L,B)} \right) \Delta Y_R,
\] (131)

while (130) is not affected.

At a given point \( \theta_M, \theta_C, \theta_R \), geometrically, the height of the polyhedron of individuals changing from occupation \( M \) to occupation \( R \) due to an increase in \( Y_R \) is given by (see Figure 5 for an illustration)

\[
\theta_R - \theta_R = \theta_M \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L,B)}{Y_R(L,B)} \right) \Delta Y_R.
\] (132)
The density at this point is
\[ f(\theta_M, \theta_R, \theta_C) = f\left(\theta_M, \theta_M \frac{Y_M(L, B)}{Y_R(L, B)}, \theta_C\right). \tag{133} \]

In order to compute the mass of individuals who switch from occupation \( M \) to \( R \) at a given \( \theta_M \) with \( \theta_R = \theta_M \frac{Y_M(L, B)}{Y_R(L, B)} \), we first need to integrate over all values \( \theta_C \) for which individuals do not work in occupation \( C \). This range of values is given by \( \frac{\partial_c}{\partial_c} \frac{Y_M(L, B)}{Y_C(L, B)} \). Integrating over the density yields
\[
\int_{\theta_C}^{\theta_M} f\left(\theta_M, \theta_M \frac{Y_M(L, B)}{Y_R(L, B)}, \theta_C\right) d\theta_C,
\]
which at a given \( \theta_M \) corresponds to a slice of the surface of indifference between sectors \( M \) and \( R \). To arrive at the first expression for the mass of switchers, we need to multiply this slice of the surface by the height of the polyhedron of switchers, \( \theta_M^R - \theta_M \), leading to
\[
\theta_M \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) \Delta Y_R \times \int_{\theta_C}^{\theta_M} f\left(\theta_M, \theta_M \frac{Y_M(L, B)}{Y_R(L, B)}, \theta_C\right) d\theta_C. \tag{135} \]

In order to compute the income moving from occupation \( M \) to occupation \( R \), due to \( \Delta Y_R \), write income as \( \theta_M Y_M(L, B) \ell(\theta_M Y_M(L, B)) \), multiply by the mass of switchers at \( \theta_M \) with \( \theta_R = \theta_M \frac{Y_M(L, B)}{Y_R(L, B)} \) and integrate over \( \theta_M \), leading to
\[
\int_{\theta_M}^{\theta_M} \theta_M Y_M(L, B) \ell\left(\theta_M Y_M(L, B)\right) \times \theta_M \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) \Delta Y_R \times \int_{\theta_C}^{\theta_M} f\left(\theta_M, \theta_M \frac{Y_M(L, B)}{Y_R(L, B)}, \theta_C\right) d\theta_C \ d\theta_M. \tag{136} \]

Next, apply the change of variables \( w = \theta_M Y_M(L, B) \), which implies \( d\theta_M = dw \frac{1}{Y_M(L, B)} \), to obtain
\[
\frac{1}{Y_M(L, B)} \int_{\frac{Y_M(L, B)}{Y_R(L, B)}}^{\frac{Y_M(L, B)}{Y_R(L, B)}} w^2 \ell(w) \times \theta_M \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) \Delta Y_R \times \int_{\theta_C}^{\theta_M} f\left(\frac{w}{Y_M(L, B)}, \frac{w}{Y_R(L, B)}, \theta_C\right) d\theta_C \ dw. \tag{137} \]

Now use that \( \Delta Y_R = Y_R(L, B) \beta_B^R(L, B) \Delta B \) to get
\[
\frac{\beta_B^R(L, B) Y_R(L, B)}{Y_M(L, B)} \int_{\frac{Y_M(L, B)}{Y_R(L, B)}}^{\frac{Y_M(L, B)}{Y_R(L, B)}} w^2 \ell(w) \times \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) \times \int_{\theta_C}^{\theta_M} f\left(\frac{w}{Y_M(L, B)}, \frac{w}{Y_R(L, B)}, \theta_C\right) d\theta_C \ dw \Delta B. \tag{138} \]

Use that \( \frac{\partial}{\partial Y_R} \left( \frac{Y_M(L, B)}{Y_R(L, B)} \right) = -\frac{Y_M(L, B)}{Y_R(L, B)^2} \) to write
\[
-\frac{\beta_B^R(L, B)}{Y_R(L, B)^2} \frac{1}{Y_M(L, B) Y_R(L, B)} \int_{\frac{Y_M(L, B)}{Y_R(L, B)}}^{\frac{Y_M(L, B)}{Y_R(L, B)}} w^2 \ell(w) \times \int_{\theta_C}^{\theta_M} f\left(\frac{w}{Y_M(L, B)}, \frac{w}{Y_R(L, B)}, \theta_C\right) d\theta_C \ dw \Delta B. \tag{139} \]
Finally, use that in order to obtain the effect on $\frac{1}{Y_M(L,B)}y(w)$ I have to divide by $Y_M(L,B)$, leading to

$$
- \beta_B^{\text{\text{\footnotesize RM}}}(L,B) \frac{1}{Y_M(L,B)^2 Y_R(L,B)} \int_{\theta_L}^{\theta_R} \frac{w^2 \ell(w)}{} \times 
\int_{\theta_L}^{\theta_R} f \left( \frac{w}{Y_M(L,B)}, \frac{w}{Y_R(L,B)}, \theta_C \right) d\theta_C \frac{d\Delta B}{dB}.
$$

(140)

Now define

$$S_{\text{\text{\footnotesize RM}}}(L,B) \equiv - \frac{1}{Y_M(L,B)^2 Y_R(L,B)} \int_{\theta_L}^{\theta_R} \frac{w^2 \ell(w)}{} \times 
\int_{\theta_L}^{\theta_R} f \left( \frac{w}{Y_M(L,B)}, \frac{w}{Y_R(L,B)}, \theta_C \right) d\theta_C \frac{d\Delta B}{dB}.
$$

(141)

In an analogue way, derive

$$S_{\text{\text{\footnotesize CM}}}(L,B) \equiv - \frac{1}{Y_M(L,B)^2 Y_C(L,B)} \int_{\theta_L}^{\theta_R} \frac{w^2 \ell(w)}{} \times 
\int_{\theta_L}^{\theta_R} f \left( \frac{w}{Y_M(L,B)}, \theta_R, \frac{w}{Y_C(L,B)} \right) d\theta_R \frac{d\Delta B}{dB}.
$$

(142)

Before providing expressions for the other terms, I repeat (part of) a Lemma from Rothschild and Scheuer (2014):

**Lemma 1.** With $I \equiv \{ M, R, C \}$, $\sum_{i \in I} C_{ij}(L,B) = \sum_{i \in I} S_{ij}(L,B) = 0$ for all $j \in I$.

By Lemma 1

$$S_{\text{\text{\footnotesize MM}}}(L,B) = - S_{\text{\text{\footnotesize RM}}}(L,B) - S_{\text{\text{\footnotesize CM}}}(L,B).
$$

(143)

This is intuitive: the inflow into occupation $M$ is equal to the flows from occupations $R$ and $C$ into $M$.

Similarly, derive

$$S_{\text{\text{\footnotesize MR}}}(L,B) \equiv - \frac{1}{Y_M(L,B)^2 Y_R(L,B)} \int_{\theta_L}^{\theta_R} \frac{w^2 \ell(w)}{} \times 
\int_{\theta_L}^{\theta_R} f \left( \frac{w}{Y_M(L,B)}, \frac{w}{Y_R(L,B)}, \theta_C \right) d\theta_C \frac{d\Delta B}{dB}.
$$

(144)

$$S_{\text{\text{\footnotesize CR}}}(L,B) \equiv - \frac{1}{Y_R(L,B)^2 Y_C(L,B)} \int_{\theta_L}^{\theta_R} \frac{w^2 \ell(w)}{} \times 
\int_{\theta_L}^{\theta_R} f \left( \theta_M, \frac{w}{Y_R(L,B)}, \frac{w}{Y_C(L,B)} \right) d\theta_M \frac{d\Delta B}{dB}.
$$

(145)

and by Lemma 1

$$S_{\text{\text{\footnotesize RR}}}(L,B) = - S_{\text{\text{\footnotesize MR}}}(L,B) - S_{\text{\text{\footnotesize CR}}}(L,B).
$$

(146)

**40** The corresponding Lemma in Rothschild and Scheuer (2014) is (the second part of) Lemma 6.
Finally, derive
\[
S_{MC}(L, B) \equiv - \frac{1}{Y_M(L, B) Y_C(L, B)^2} \int \frac{w^2 \ell(w)}{w L} \partial_B w, \tag{147}
\]
\[
S_{RC}(L, B) \equiv - \frac{1}{Y_R(L, B) Y_C(L, B)^2} \int \frac{w^2 \ell(w)}{w L} \partial_B w, \tag{148}
\]
and
\[
S_{CC}(L, B) = -S_{MC}(L, B) - S_{RC}(L, B). \tag{149}
\]
Having derived all occupational shift effects, it remains to combine them. The occupational shift effect which corresponds to the consistency condition for occupation \( j \) is
\[
- \mu \xi_j \sum_{i \in I} \beta_B^j(L, B) S_{ij} dB. \tag{150}
\]

**Putting everything together.** Combining the terms derived above, we get
\[
\frac{\partial \mathcal{L}(L, B)}{\partial B} = \\
\mu \left[ Y_B(L, B) - q - \sum_{i \in I} \beta_B^j(L, B) \left( I_i(L, B) + \sum_{j \in I} \xi_j (C_{ij}(L, B) + S_{ij}(L, B)) \right) \right]. \tag{151}
\]
Now use that
\[
Y_B(L, B) = (1 + \tau) q \Leftrightarrow Y_B(L, B) - q = \tau q. \tag{152}
\]
Since at the optimum \( \partial \mathcal{L}(L, B) / \partial B = 0 \), we get
\[
\tau q = \sum_{i \in I} \beta_B^j(L, B) \left( I_i(L, B) + \sum_{j \in I} \xi_j (C_{ij}(L, B) + S_{ij}(L, B)) \right). \tag{153}
\]
To further rewrite the expression, first focus on
\[
\beta_B^R(L, B) \left( I_R(L, B) + \sum_{j \in I} \xi_j (C_{Rj}(L, B) + S_{Rj}(L, B)) \right)
\]
\[
= - \beta_B^B(L, B) \left[ I_M(L, B) + I_C(L, B) \right. \\
- \xi_M (C_{RM}(L, B) + S_{RM}(L, B)) \\
- \xi_R (C_{RR}(L, B) + S_{RR}(L, B)) \\
- \xi_C (C_{RC}(L, B) + S_{RC}(L, B)) \left. \right] 	ag{154}
\]
\[
= - \beta_B^B(L, B) \left[ I_M(L, B) + I_C(L, B) ight. \\
+ \xi_M (C_{MM}(L, B) + C_{CM}(L, B) + S_{MM}(L, B) + S_{CM}(L, B)) \\
+ \xi_R (C_{MR}(L, B) + C_{CR}(L, B) + S_{MR}(L, B) + S_{CR}(L, B)) \\
+ \xi_C (C_{MC}(L, B) + C_{CC}(L, B) + S_{MC}(L, B) + S_{CC}(L, B)) \left. \right],
\]
where the first step uses \( \sum_{i \in I} I_i = \sum_{i \in I} R_i = 0 \) and the second step uses Lemma 1. Substituting (154) for the respective expression in (153) and collecting terms yields

\[
\tau_q = \left( \beta_C^B (\mathbf{L}, B) - \beta_R^B (\mathbf{L}, B) \right) \left( I_C (\mathbf{L}, B) + \sum_{j \in I} \xi_j (C_{Cj} (\mathbf{L}, B) + S_{Cj} (\mathbf{L}, B)) \right) \\
+ \left( \beta_M^B (\mathbf{L}, B) - \beta_R^B (\mathbf{L}, B) \right) \left( I_M (\mathbf{L}, B) + \sum_{j \in I} \xi_j (C_{Mj} (\mathbf{L}, B) + S_{Mj} (\mathbf{L}, B)) \right).
\]

Finally, use that

\[
\varepsilon_{Y_C/Y_R,B} (\mathbf{L}, B) = \frac{\partial (Y_C (\mathbf{L}, B) / Y_R (\mathbf{L}, B))}{\partial B} \frac{B}{Y_C (\mathbf{L}, B) / Y_R (\mathbf{L}, B)} \\
= B \left( \beta_C^B (\mathbf{L}, B) - \beta_R^B (\mathbf{L}, B) \right),
\]

and

\[
\varepsilon_{Y_M/Y_R,B} (\mathbf{L}, B) = \frac{\partial (Y_M (\mathbf{L}, B) / Y_R (\mathbf{L}, B))}{\partial B} \frac{B}{Y_M (\mathbf{L}, B) / Y_R (\mathbf{L}, B)} \\
= B \left( \beta_M^B (\mathbf{L}, B) - \beta_R^B (\mathbf{L}, B) \right),
\]

to arrive at

\[
\tau q B = \varepsilon_{Y_C/Y_R,B} (\mathbf{L}, B) \left( I_C (\mathbf{L}, B) + \sum_{j \in I} \xi_j (C_{Cj} (\mathbf{L}, B) + S_{Cj} (\mathbf{L}, B)) \right) \\
+ \varepsilon_{Y_M/Y_R,B} (\mathbf{L}, B) \left( I_M (\mathbf{L}, B) + \sum_{j \in I} \xi_j (C_{Mj} (\mathbf{L}, B) + S_{Mj} (\mathbf{L}, B)) \right).
\]

C Log-Likelihood function

Let individual observations be indexed by \( j = 1 \ldots N \). I define the following indicator function

\[
\mathbb{I}_i^j = \begin{cases} 
1 & \text{if } j \text{ works in occupation } i \\
0 & \text{otherwise},
\end{cases}
\]

where \( i \in \{M, R, C\} \). Moreover, denote by \( Y_i \) the skill-price in occupation \( i \), and by \( w^j \) the wage of individual \( j \). \( \sigma_i, \rho_{MR}, \rho_{MC}, \rho_{RC} \) are parameters of the tri-variate Normal distribution which are to be estimated. The log-likelihood function is then
\[ L = \sum_{j=1}^{N} I \left[ \ln \sigma_M + \ln \phi \left( \frac{\ln w_j - Y_M}{\sigma_M} \right) + \ln \Phi \left( \frac{\ln w_j - Y_R}{\sigma_R} - \rho_{MR} \frac{\ln w_j - Y_M}{\sigma_M} \sqrt{1 - \rho_{MR}^2}, \frac{\ln w_j - Y_M}{\sigma_M} - \rho_{MC} \frac{\ln w_j - Y_M}{\sigma_M} ; \rho_{RC,M} \right) \right] \]

\[ + \ln \Phi \left( \frac{\ln w_j - Y_R}{\sigma_R} - \rho_{MR} \frac{\ln w_j - Y_R}{\sigma_R} \sqrt{1 - \rho_{MR}^2}, \frac{\ln w_j - Y_R}{\sigma_R} - \rho_{RC} \frac{\ln w_j - Y_R}{\sigma_R} ; \rho_{MC,R} \right) \]

\[ + \ln \Phi \left( \frac{\ln w_j - Y_C}{\sigma_C} - \rho_{MC} \frac{\ln w_j - Y_C}{\sigma_C} \sqrt{1 - \rho_{MC}^2}, \frac{\ln w_j - Y_C}{\sigma_C} - \rho_{RC} \frac{\ln w_j - Y_C}{\sigma_C} ; \rho_{MR,C} \right) \]

Here \( \phi \) is the standard Normal density and \( \Phi \) is the CDF of a bivariate standard Normal distribution with covariance \( \rho_{ab,c} \). Following Bi and Mukherjea (2010), we have

\[ \rho_{RC,M} \equiv \frac{\rho_{RC} - \rho_{MR} \rho_{MC}}{\sqrt{1 - \rho_{MR}^2} \sqrt{1 - \rho_{MC}^2}}, \rho_{MC,R} \equiv \frac{\rho_{MC} - \rho_{MR} \rho_{RC}}{\sqrt{1 - \rho_{MR}^2} \sqrt{1 - \rho_{RC}^2}}, \rho_{MR,C} \equiv \frac{\rho_{MR} - \rho_{MC} \rho_{RC}}{\sqrt{1 - \rho_{MC}^2} \sqrt{1 - \rho_{RC}^2}}. \]
D Robustness

Table 7: Robustness for different targeted elasticities $\varepsilon^{GE}_{\pi R,B}$

<table>
<thead>
<tr>
<th>Targeted elasticity</th>
<th>10 x higher</th>
<th>20 x higher</th>
<th>30 x higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occup. Choice</td>
<td>No, Yes</td>
<td>No, Yes</td>
<td>No, Yes</td>
</tr>
<tr>
<td>Robot Tax</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.84</td>
<td>0.41</td>
<td>3.56</td>
<td>0.39</td>
</tr>
<tr>
<td>3.25</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Welf. impact of robot tax per cap. in $</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.88</td>
<td>0.11</td>
<td>14.54</td>
<td>0.22</td>
</tr>
<tr>
<td>19.86</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elasticities</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon^{GE}_{Y_M/Y_R,B}$</td>
<td>0.0534</td>
<td>0.0301</td>
<td>0.1106</td>
</tr>
<tr>
<td>$\varepsilon^{GE}_{Y_C/Y_R,B}$</td>
<td>0.0534</td>
<td>0.0511</td>
<td>0.1106</td>
</tr>
<tr>
<td>$\varepsilon^{GE}_{\pi_M,B}$</td>
<td>0.0077</td>
<td>0.0075</td>
<td>0.0172</td>
</tr>
<tr>
<td>$\varepsilon^{GE}_{\pi_R,B}$</td>
<td>-0.0180</td>
<td>-0.0121</td>
<td>-0.0360</td>
</tr>
<tr>
<td>$\varepsilon^{GE}_{\pi_C,B}$</td>
<td>0.0353</td>
<td>0.0015</td>
<td>0.0744</td>
</tr>
</tbody>
</table>

Production Function Parameters

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>30.53</th>
<th>30.26</th>
<th>29.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.43</td>
<td>0.41</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.99</td>
<td>0.98</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>5.36</td>
<td>5.97</td>
<td>6.73</td>
<td></td>
</tr>
</tbody>
</table>

Note: Values are obtained by calibrating the model to an elasticity $\varepsilon^{GE}_{\pi R,B}$ which is, respectively, 10, 20, and 30 times higher than the elasticity in the baseline calibration. The targeted elasticity is indicated in bold. The corresponding production function parameters are reported as well as the implied elasticities, both for a model without and with occupational choice. The tax and welfare computations are based on the same procedure as for the baseline model and are also conducted for a model without and with occupational choice. The welfare impact is expressed in 2016 dollar values per capita per year.

E Data Appendix

I obtain data on wages and occupational choice from the Current Population Survey (CPS) Merged Outgoing Rotation Groups (MORG) as prepared by the National Bureau of Economic Research (NBER). The data cover the years from 1979 to 2016. To make results comparable with Acemoglu and Restrepo (2017), I focus on data from 1993 which is the first in year in which data on industrial robots is available for the US.

41see http://www.nber.org/data/morg.html
Table 8: Occupation classification from Acemoglu and Autor (2011)

<table>
<thead>
<tr>
<th>Recode</th>
<th>Description</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Managers</td>
<td>Abstract</td>
</tr>
<tr>
<td>12</td>
<td>Professionals</td>
<td>Abstract</td>
</tr>
<tr>
<td>13</td>
<td>Technicians</td>
<td>Abstract</td>
</tr>
<tr>
<td>21</td>
<td>Sales</td>
<td>Routine cogn.</td>
</tr>
<tr>
<td>22</td>
<td>Office and admin</td>
<td>Routine cogn.</td>
</tr>
<tr>
<td>23</td>
<td>Production, craft and repair</td>
<td>Routine man.</td>
</tr>
<tr>
<td>31</td>
<td>Operators, fabricators and laborers</td>
<td>Routine man.</td>
</tr>
<tr>
<td>32</td>
<td>Protective service</td>
<td>Services</td>
</tr>
<tr>
<td>33</td>
<td>Food prep, buildings and grounds, cleaning</td>
<td>Services</td>
</tr>
<tr>
<td>34</td>
<td>Personal care and personal services</td>
<td>Services</td>
</tr>
</tbody>
</table>

E.1 Sample

Selection of the sample follows Acemoglu and Autor (2011). I include individuals aged 16 to 64 whose usual weekly hours worked exceed 35. Hourly wages are obtained by dividing weekly earnings by usual hours worked. All wages are converted into 2016 dollar values using the personal consumption expenditures chain-type price index.\textsuperscript{42} The highest earnings in the CPS are top-coded. I therefore winsorize earnings by multiplying top-coded earnings by 1.5. Like Acemoglu and Autor (2011), I exclude those individuals who earn less than 50% of the 1982 minimum wage (3.35$) converted to 2016-dollars. Self-employed individuals are excluded, as are individuals whose occupation does not have an occ1990dd classification. Like Acemoglu and Autor (2011), I exclude individuals employed by the military as well as agricultural occupations. As will be discussed in Section E.2, I also exclude the following occupations: Police, detectives and private investigators, Fire fighting, prevention and inspection, Other law enforcement: sheriffs, bailiffs, correctional institution officers. Observations are weighted by CPS sample weights.

E.2 Classifying occupations

I classify occupations into three categories: manual non-routine, manual routine and cognitive. To do so, I proceed in several steps.

Two-digit classification as in Acemoglu and Autor (2011). First, I apply the classification from David Dorn.\textsuperscript{43} Next, I follow Acemoglu and Autor (2011) and group occupations into the following categories: Managers, Professionals, Technicians, Sales, Office and admin, Personal care and personal services, Protective service, Food prep, buildings and grounds, cleaning, Agriculture, Production, craft and repair, Operators, fabricators and laborers. Autor and Dorn (2013) highlight that protective services is a heterogeneous category with wages in police, firefighters and other law enforcement occupations being substantially higher than in other protective services occupations. I therefore exclude these occupations from the analysis. Figure 6 shows a box-plot of log wages for the different recoded occupations, ordered by median wage.

Classification into three categories. I base my classification of occupations on Acemoglu and Autor (2011) who use two types of classifications. The first classification with four categories

\textsuperscript{42} I obtain the price index from https://fred.stlouisfed.org/series/DPCERG3A086NBEA
\textsuperscript{43} See http://www.ddorn.net/data.htm
distinguishes between abstract, routine cognitive, routine manual and services occupations, as described in Table 8.\textsuperscript{44} A second classification with three categories groups routine cognitive and routine manual occupations together.\textsuperscript{45} In the paper, I follow the latter classification.\textsuperscript{46} Figure 7 illustrates the distribution of log hourly wages for this classification.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{log_wage_distribution.png}
\caption{Distribution of log wages for 2-digit occupations}
\textit{Note:} [Source] NBER CPS Merged Outgoing Rotation Groups. The sample is explained in Section E.1. Data is for the year 1993. Wages are in 2016-Dollars.
\end{figure}

\textsuperscript{44}The classification is described in tab-reg-wages-census-table-specs.do

\textsuperscript{45}While four categories might capture the impact of robots in a more nuanced way, I choose to avoid having to estimate a four-dimensional skill distribution, since estimating a three-dimensional skill-distribution is already challenging.

\textsuperscript{46}While one might argue that robots primarily substitute for manual routine workers, Acemoglu and Restrepo (2017) find negative effects of robot exposure on employment also for occupations which can be classified as cognitive routine.
Figure 7: Distribution of log wages for the three categories of occupations

*Note:* [Note] NBER CPS Merged Outgoing Rotation Groups. The sample is explained in Section E.1. The classification into three categories is explained in Section E.2. Data is for the year 1993. Wages are in 2016-Dollars.
References


